

17409
SEARCH REQUEST FORM

Examiner # (Mandatory): _____ Requester's Full Name: _____

Art Unit _____ Location (Bldg/Room#): 7E09 Phone (circle 305 306 308) _____

Serial Number: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms): _____

Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

STAFF USE ONLYSearcher: PROB

Searcher Phone #: _____

Searcher Location: _____

Date Picked Up: 8-3Date Completed: 8-4-99Clerical Prep Time: 6 9Terminal Time: 6

Number of Databases: _____

Type of Search☒ N.A. Sequence☒ A.A. Sequence☐ Structure (#)☐ Bibliographic☐ Litigation1☐ Fulltext☐ Procurement☐ Other**Vendors (include cost where applicable)**☐ STN☐ Questel/Orbit☐ Lexis/Nexis☐ WWW/Internet☒ MP3 In-house sequence systems (list)☐ Dialog☐ Dr. Link☐ Westlaw☐ Other (specify)

This Page Blank (uspto)

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840.

(W.T.)

| MPsrch_nn | n.a. | - | n.a. | database search, using Smith-Waterman algorithm |
|-----------|------|---|------|---|
| 1 | 1 | 1 | 1 | 1 |

```
Run on: Tue Aug 3 21:17:35 1999; MasPar time 3626.14 Seconds
1536 403 million coll updates/sec
```

Tabular output not generated.

| | |
|--------------|------------------------------|
| Title: | >US-09-240-410-1 |
| Description: | (1-2010) from US09240410.seq |

N.A. Sequence: 1 ATGACGCTCTCCGCGCCG...TGGTCCACTAGGCGCTCCCG 2010

Comp:

Scoring table: TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 646147 seqs, 1385953633 bases x 2

Post-processing: Minimum Match 0%

Database:

```

1:em_ba1 2:em_ba2 3:em_fun 4:em_htg 5:em_hun1 6:em_hun2
7:em_in 8:em_com 9:em_or 10:em_ov 11:em_pat 12:em_ph
13:em_p1 14:em_ro 15:em_sts 16:em_v1
genbank111
Database:

```

Database:

Statistics: Mean 11.468; Variance 5.912; scale 1.940

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------|------------------------|------------|
| 1 | 2010 | 100.0 | 2594 | 31 | AF071543 | Homo sapiens semaphori | 0.00e+00 |
| 2 | 2010 | 100.0 | 2661 | 31 | AF065492 | Homo sapiens GPI-anch | 0.00e+00 |
| 3 | 2004 | 99.7 | 2636 | 31 | AF030698 | Homo sapiens semaphori | 0.00e+00 |
| 4 | 1444 | 71.8 | 2608 | 32 | AF017532 | Mus musculus mRNA for | 0.00e+00 |
| 5 | 847 | 42.1 | 1192 | 32 | AF030699 | Mus musculus semaphori | 0.00e+00 |
| 6 | 375 | 18.7 | 8861 | 31 | AF018647 | Homo sapiens semaphori | 9.69e-2707 |
| 7 | 131 | 6.5 | 1874 | 37 | AHU18243 | Alcelaphine herpesvitu | 7.14e-74 |
| 8 | 131 | 6.5 | 13608 | 37 | AF005370 | Alcelaphine herpesvitu | 7.14e-74 |
| 9 | 42 | 2.1 | 7218 | 25 | 166494 | Sequence 14 from paten | 2.65e-09 |
| 10 | 38 | 1.9 | 965 | 25 | AF024229 | Sequence 22 from paten | 7.67e-07 |
| 11 | 38 | 1.9 | 1056 | 32 | MU87256 | Mustela vison GR dnuc | 7.67e-07 |
| 12 | 38 | 1.9 | 2582 | 32 | MUSICP2RA | Mus musculus insulin-1 | 7.67e-07 |
| 13 | 38 | 1.9 | 7218 | 25 | 166494 | Sequence 14 from paten | 7.67e-07 |

| | | | | | | | | | |
|---|----|----|-----|--------|----|-----------|---------------------------|----------|----------|
| C | 14 | 38 | 1.9 | 8810 | 32 | RNU59809 | Rattus norvegicus | man | 7.67e-07 |
| | 15 | 37 | 1.8 | 215 | 25 | 128278 | Sequence 5 from patent | 3.05e-06 | |
| | 16 | 37 | 1.8 | 215 | 25 | 128278 | Sequence 5 from patent | 3.05e-06 | |
| C | 17 | 36 | 1.8 | 965 | 25 | AR024229 | Sequence 22 from patent | 1.20e-05 | |
| | 18 | 36 | 1.8 | 1056 | 23 | MW087256 | Mustela vison GT dinuc | 1.20e-05 | |
| C | 19 | 37 | 1.8 | 1663 | 23 | MW092534 | Mustela vison microsat | 3.05e-06 | |
| | 20 | 37 | 1.8 | 2336 | 24 | AF022946 | Gallus gallus microsat | 3.05e-06 | |
| | 21 | 36 | 1.8 | 2919 | 31 | U528369 | Homo sapiens semaphorin | 1.20e-05 | |
| C | 22 | 36 | 1.8 | 8877 | 32 | MM004710 | Homo sapiens semaphorin | 1.20e-05 | |
| | 23 | 36 | 1.8 | 36534 | 30 | U73167 | Mus musculus domestica | 1.20e-05 | |
| C | 24 | 36 | 1.8 | 36552 | 30 | AC005094 | Homo sapiens cosmid cl | 1.20e-05 | |
| | 25 | 36 | 1.8 | 123612 | 31 | AC006208 | Human chromosome 3p21.1 | 1.20e-05 | |
| C | 26 | 35 | 1.7 | 74371 | 31 | AC005369 | Homo sapiens chromosome | 4.55e-05 | |
| | 27 | 35 | 1.7 | 74371 | 31 | AC005369 | Homo sapiens chromosome | 4.55e-05 | |
| C | 28 | 32 | 1.6 | 201 | 25 | A10158 | Homo sapiens gene for pre | 2.43e-03 | |
| | 29 | 32 | 1.6 | 201 | 25 | A10158 | Synthetic DNA for prep | 2.43e-03 | |
| C | 30 | 32 | 1.6 | 201 | 25 | A10161 | Synthetic DNA for prep | 2.43e-03 | |
| | 31 | 32 | 1.6 | 461 | 32 | MSINSUL01 | Mouse insulin-like gene | 2.43e-03 | |
| C | 32 | 32 | 1.6 | 879 | 30 | U52989 | Homo sapiens clone D11 | 2.43e-03 | |
| | 33 | 32 | 1.6 | 1270 | 32 | AF020421 | Rat alpha-2-macroglobu | 2.43e-03 | |
| C | 34 | 32 | 1.6 | 1864 | 30 | AF020276 | Homo sapiens spinocere | 2.43e-03 | |
| | 35 | 32 | 1.6 | 2040 | 23 | CF029765 | Cantis familiaris alpha | 2.43e-03 | |
| C | 36 | 32 | 1.6 | 2265 | 32 | AF080091 | Mus musculus semaphorin | 6.62e-04 | |
| | 37 | 33 | 1.6 | 2383 | 32 | AF080090 | Mus musculus semaphorin | 6.62e-04 | |
| C | 38 | 32 | 1.6 | 3367 | 31 | AF032105 | Homo sapiens semaphorin | 2.43e-03 | |
| | 39 | 33 | 1.6 | 3518 | 32 | RATNT18A | Homo sapiens ataxin-7 | 6.62e-04 | |
| C | 40 | 32 | 1.6 | 3819 | 31 | AF106861 | Rattus norvegicus neur | 2.43e-03 | |
| | 41 | 32 | 1.6 | 3969 | 30 | H5SCA7 | Homo sapiens attractin | 2.43e-03 | |
| C | 42 | 32 | 1.6 | 4029 | 32 | MM026348 | Homo sapiens mRNA for | 2.43e-03 | |
| | 43 | 32 | 1.6 | 5633 | 29 | BSU17033 | Mus musculus insulin-1 | 2.43e-03 | |
| C | 44 | 33 | 1.6 | 7913 | 23 | HCWCMR | Human 180 kDa transmem | 6.62e-04 | |
| | 45 | 32 | 1.6 | 8670 | 27 | CRAKRG7 | Bovine cation-independent | 2.43e-03 | |
| | | | | | | | Chlamydomonas reinhard | | |

ALIGNMENTS

| LOCUS | 1 | RESULT |
|------------|---|------------|
| AF071542 | 2594 bp | mRNA |
| DEFINITION | Homo sapiens semaphorin K1 mRNA, complete cds. | PRI |
| AF071542 | | |
| ACCESSION | g3930578 | |
| NID | | |
| VERSION | AF071542.1 | GI:3930578 |
| KEYWORDS | | |
| SOURCE | human. | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | |
| AUTHORS | Primates; Catarrhini; Hominae; Homo. | |
| TITLE | 1 (bases 1 to 2594) | |
| JOURNAL | Xu,X., Ng,S., Wu,Z.L., Nguyen,D., Homburger,S., Seidel-Dugan,C., | |
| MEDLINE | Ebens,A. and Luo,Y. | |
| REFERENCE | Human semaphorin K1 is glycosylphosphatidylinositol-linked and | |
| AUTHORS | defines a new subfamily of viral-related semaphorins | |
| TITLE | J. Biol. Chem. 273 (35), 22428-22434 (1998) | |
| REFERENCE | 98380463 | |
| AUTHORS | 2 (bases 1 to 2594) | |
| TITLE | Xu,X. and Luo,Y. | |
| REFERENCE | Direct Submission | |
| AUTHORS | Submitted (10-JUN-1998) Biology, Exelixis Pharmaceuticals, Inc., 2600 | |
| TITLE | Littlefield Ave, South San Francisco, CA 94080, USA | |
| REFERENCE | 3 (bases 1 to 2594) | |
| AUTHORS | Xu,X. and Luo,Y. | |
| TITLE | Direct Submission | |
| REFERENCE | Submitted (30-NOV-1998) Biology, Exelixis Pharmaceuticals, Inc., 2600 | |
| AUTHORS | Littlefield Ave, South San Francisco, CA 94080, USA | |
| TITLE | Sequence update by submitter | |
| REFERENCE | On Dec 1, 1998 this sequence version replaced gi:3493630. | |
| REMARK | | |
| COMMENT | | |
| FEATURES | | |
| SOURCE | | |

1. .2594

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="15"

/map="15q22"

1. .2001

August 78

NOT prior

include with

```

/note="sema K1: glycosylphosphatidylinositol-linked
protein; similar to Alcelaphine herpesvirus 1 semaphorin"
/codon_start=1
/product="semaphorin K1"
/protein_id="AAC80456.1"
/db_xref="PID:93930579"
/db_xref="GI:3930579"
/translation="MTPPGGRAASAPARVPAPRIGLRILLILMAAASNS
OGHRSGRPRITAAVKGHGQDPDPCGTEPRITYLVFHEGGSSVYAGSGKTYLDFPE
GKNSVRYTNGSTKSGSLDRDDENTITLERSSEGLAGIARHPSCWNLVNGTV
VPLDEMRATFPSPDENSLVLFEGDEVSTITRKEQENKIPREFRRIGESLSYSDTV
MONEQIKTAVIHODQADDKIYVFERIDANDKNKAPLANTYRAOLCRGOGGESSIT
SVSWNTEFLKMLVCSDAITNKNFNRLDDVLLDPSQOMDITVYGVFSNMYASAV
CVYSIGIDIKVETSLKGYHSSLPDGGCLDDPDIETFOVARDHEVORARE
PMGELKTPILFHSKTHYOKVAHQAASIGETFHLVLTIDGTIHKHYVPEQHSRA
NMEIIOFERAAAIQMTSLDAERRKATLYVSSOMVSOVPLDICVGGCGCLMSND
PYCGMDGRCSTIYSERSVYQSTINPAPHECNPAPDKAPLDKVSILAPSRITLSC
PMERHATTSRKHEENVEQSCPEGHQSNCLFLFENLTADQYGFCEAQEGSYFREDA
QHWHLLEDGDGMALHILGHACALASMLGVLPLNTLGLLVH"
514 a 814 c 756 g 510 t
BASE COUNT
ORIGIN

```

| | | | |
|--------------------------------|------------------|---|---------------------|
| Query Match: 100.0%; | | Score: 20.10; DB: 31; | Length: 2594; |
| Best Local Similarity: 100.0%; | | Pred. NO: 0.00e+00; | |
| Matches: 2010; | Conservative: 0; | Mismatches: 0; | Indels: 0; Gaps: 0; |
| D | 1 | ATGACGCTCTCCGCCGAGCTGCGCCGCCACCGCCACCGCGCGCCCGCTCCCTGGC | 60 |
| Q | 1 | ATGACGCTCTCCGCCGAGCTGCGCCGCCACCGCCACCGCGCGCCCGCTCCCTGGC | 60 |
| D | 61 | CCGCGGCTCGGTTGGGGCTTCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCGGGCGCC | 120 |
| Q | 61 | CCGCGGCTCGGTTGGGGCTTCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCGGGCGCC | 120 |
| D | 121 | GCCGCTCCGCCGAGGGGCACTTAAGAGGGAGCCCGGCAATTCGCGCTGGAAAGC | 180 |
| Q | 121 | GCCGCTCCGCCGAGGGGCACTTAAGAGGGAGCCCGGCACTTCGCGCTGGAAAGC | 180 |
| D | 181 | CATTAAGGGCAGSACCGGGTGGACTTTGGCGAGCTAGCGCGCACAGGTCGTTTTCAC | 240 |
| Q | 181 | CATTAAGGGCAGSACCGGGTGGACTTTGGCGAGCTAGCGCGCACAGGTCGTTTTCAC | 240 |
| D | 241 | GAGCCAGGCACTCTCTGTGTGGGTGGAGAGAGCTGCAAGCTTCACTCTTTGACTTC | 300 |
| Q | 241 | GAGCCAGGCACTCTCTGTGTGGGTGGAGAGAGCTGCAAGGCTCACTCTTTGACTTC | 300 |
| D | 301 | CCGAGGGCAGAAACGATCTGTGCGAGCGTGAATTCGGCTCCCAAGGGGCTCTGT | 360 |
| Q | 301 | CCGAGGGCAGAAACGATCTGTGCGAGCGTGAATTCGGCTCCCAAGGGGCTCTGT | 360 |
| D | 361 | CTGATTAAGCGGGACTCGAGAACTACATCACTCTCTGAGAGGGGAGTAGGGGCTG | 420 |
| Q | 361 | CTGATTAAGCGGGACTCGAGAACTACATCACTCTCTGAGAGGGGAGTAGGGGCTG | 420 |
| D | 421 | CTGGCTCTGTGGCACAAACGCCCGGACCCAGCTGCTGGAAGCTGGTAATGCACTGTG | 480 |
| Q | 421 | CTGGCTCTGTGGCACAAACGCCCGGACCCAGCTGCTGGAAGCTGGTAATGCACTGTG | 480 |
| D | 481 | GTCGCAATTTGGCGAGTAGAGGCTAGGCCCTTAGCCCGGACAGAACTCCCTGGTT | 540 |
| Q | 481 | GTCGCAATTTGGCGAGTAGAGGCTAGGCCCTTAGCCCGGACAGAACTCCCTGGTT | 540 |
| D | 541 | CTGTTTGAAGGGGACGAGTGTATTCCACCATCCGGAAGAGGAATATCAATGGGAAGATC | 600 |
| Q | 541 | CTGTTTGAAGGGGACGAGTGTATTCCACCATCCGGAAGAGGAATATCAATGGGAAGATC | 600 |
| D | 601 | CTTGCGTCCGCGCATCCGGGGCGAGAGTGAAGCTGTACACCAAGTATCTGTATGACG | 660 |
| Q | 601 | CTTGCGTCCGCGCATCCGGGGCGAGAGTGAAGCTGTACACCAAGTATCTGTATGACG | 660 |
| D | 661 | AACCCACAGTTATCAAAAGCCACATCTGTGACCAAGACAGGCTTACATGACAAGATC | 720 |
| Q | 661 | AACCCACAGTTATCAAAAGCCACATCTGTGACCAAGACAGGCTTACATGACAAGATC | 720 |
| D | 720 | ATGACGCTCTCCGCCGAGCTGCGCCGCCACCGCCACCGCGCGCCCGCTCCCTGGC | 780 |
| Q | 720 | ATGACGCTCTCCGCCGAGCTGCGCCGCCACCGCCACCGCGCGCCCGCTCCCTGGC | 780 |

| | | | |
|---|------|--|------|
| D | 721 | TACTACTCTTCCGAGAGACAACTCTGACAAAGATCTCTAAGGCTCTCTCAATGTGTCC | 780 |
| Q | 721 | TACTACTCTTCCGAGAGACAACTCTGACAAAGATCTCTAAGGCTCTCTCAATGTGTCC | 780 |
| D | 781 | CGTGTGGCCCAATTTGTGCAGGGGGGACAGAGGTGGGGAAGTTCACTGTAGCTCCAG | 840 |
| Q | 781 | CGTGTGGCCCAATTTGTGCAGGGGGGACAGAGGTGGGGAAGTTCACTGTAGCTCTCAAG | 840 |
| D | 841 | TGGAACTCTTCTCAAAAGCATGTGTATGACGTATGTCTGCCAACAAGACTTC | 900 |
| Q | 841 | TGGAACTCTTCTCAAAAGCATGTGTATGACGTATGTCTGCCAACAAGACTTC | 900 |
| D | 901 | AACAGGCTGCAAGAGGTTCCTGCTCCCTAGCCCCAGGGCCAGAGGAGGACACAGG | 960 |
| Q | 901 | AACAGGCTGCAAGAGGTTCCTGCTCCCTAGCCCCAGGGCCAGAGGAGGACACAGG | 960 |
| D | 961 | GTCTATGGTGTTCCTCCAAACCCCTGGAATCTCAGCCGCTGTGTATTCCTCGGT | 1020 |
| Q | 961 | GTCTATGGTGTTCCTCCAAACCCCTGGAATCTCAGCCGCTGTGTATTCCTCGGT | 1020 |
| D | 1021 | GACATTGACAAGAGTTCCTGCTACCTCCACTCAAGGGCTACACATCAGACCTTCCCAAC | 1080 |
| Q | 1021 | GACATTGACAAGAGTTCCTGCTACCTCCACTCAAGGGCTACACATCAGACCTTCCCAAC | 1080 |
| D | 1081 | CCGGGCGCTGGCAAGTGCCTCCAGACAGACAGCCGATATCCACAGAGACTTCCAGAG | 1140 |
| Q | 1081 | CCGGGCGCTGGCAAGTGCCTCCAGACAGACAGCCGATATCCACAGAGACTTCCAGAG | 1140 |
| D | 1141 | GTGAGCGCTCAACCAAGAGGTGGCCACAGAGGTGAGAGCCCATGGGGCTTGAAGAGCCA | 1200 |
| Q | 1141 | GTGAGCGCTCAACCAAGAGGTGGCCACAGAGGTGAGAGCCCATGGGGCTTGAAGAGCCA | 1200 |
| D | 1201 | TTGTTCACACTTAATATACATACACAGAAAGTGGCGCTCACCGCATGGAACCGACAC | 1260 |
| Q | 1201 | TTGTTCACACTTAATATACATACACAGAAAGTGGCGCTCACCGCATGGAACCGACAC | 1260 |
| D | 1261 | GGGGAGACCTTCATGTGTCTTACCTTACTACAGACAGGGGCACTATCCAAAGGTGGTG | 1320 |
| Q | 1261 | GGGGAGACCTTCATGTGTCTTACCTTACTACAGACAGGGGCACTATCCAAAGGTGGTG | 1320 |
| D | 1321 | GAACCGGGGAGAGAGAGACAGAGTTGGCTTAACATCATGGAAGATCCAGGCCCTTCGCG | 1380 |
| Q | 1321 | GAACCGGGGAGAGAGAGACAGAGTTGGCTTAACATCATGGAAGATCCAGGCCCTTCGCG | 1380 |
| D | 1381 | CGCGGCGCTGCCATCCAGACCATGTGGCTGATGTGACGGAGAAAGCTGTATGTAGC | 1440 |
| Q | 1381 | CGCGGCGCTGCCATCCAGACCATGTGGCTGATGTGACGGAGAAAGCTGTATGTAGC | 1440 |
| D | 1441 | TCCAGTGGGAGGTGAGACACAGGTGGCCCTGAGACCTGTGAGGTCTATGAGGGGGGCTGC | 1500 |
| Q | 1441 | TCCAGTGGGAGGTGAGACACAGGTGGCCCTGAGACCTGTGAGGTCTATGAGGGGGGCTGC | 1500 |
| D | 1501 | CACGGTTGCTCATGTCCGAGACCCCTACTGTGGGTGGGACCAAGGCGCTGATCTCC | 1560 |
| Q | 1501 | CACGGTTGCTCATGTCCGAGACCCCTACTGTGGGTGGGACCAAGGCGCTGATCTCC | 1560 |
| D | 1561 | ATCTACAGCTCCGAAGCGTAGGTGTCATCATTAATCCAGCCGAGGCCACACAAGAG | 1620 |
| Q | 1561 | ATCTACAGCTCCGAAGCGTAGGTGTCATCATTAATCCAGCCGAGGCCACACAAGAG | 1620 |
| D | 1621 | TGTCCCAACCCCAACCAAGAGCCCACTGACAGAGAGTTTCCCTGGCCCCAAACTCT | 1680 |
| Q | 1621 | TGTCCCAACCCCAACCAAGAGCCCACTGACAGAGAGTTTCCCTGGCCCCAAACTCT | 1680 |
| D | 1681 | CGCTACTACCTGAGCTGCCCATGGAATCCCGCACGCACTACTCATGGCGGCACAG | 1740 |
| Q | 1681 | CGCTACTACCTGAGCTGCCCATGGAATCCCGCACGCACTACTCATGGCGGCACAG | 1740 |
| D | 1741 | GAGAACTGGAGAGAGCTGCGAACTGGTCTACAGAGCCCACTGATCTGTTCATC | 1800 |
| Q | 1741 | GAGAACTGGAGAGAGCTGCGAACTGGTCTACAGAGCCCACTGATCTGTTCATC | 1800 |

[illegible]

priority
back to 1/92

price + mo

[illegible]

|||||
QY 1 ATGAGCCTCTCCGCCGAGCTGCCGCCACAGCAGCCGCGCTCCCTGGC 60
Db 73 CCGCGGCTCGGTGGGGCTTCCGCTGGGCTGGGGCTGCTGCTCTGGGGGCC 132
QY 61 CCGCGGCTCGGTGGGGCTTCCGCTGGGCTGGGGCTGCTGCTGGGGGCC 120
Db 133 GCGCGCTCCGCCAGGGGCTTAAGAGAGGAGCCGCCATCTTGGCGCTGGAAAAGC 192
QY 121 GCGCGCTCCGCCAGGGGCTTAAGAGAGGAGCCGCCATCTTGGCGCTGGAAAAGC 180
Db 193 CATGTAGGCGAGAGCCGGGTGGACTTTGGCCAGATGAGCCGACAGCGCTTTTCCAC 252
QY 181 CATGTAGGCGAGAGCCGGGTGGACTTTGGCCAGATGAGCCGACAGCGCTTTTCCAC 240
Db 253 GAGCCAGGAGCTCTCTGTGTGGTGGAGAGCTGGCAAGGTCTACTCTTTGACTTC 312
QY 241 GAGCCAGGAGCTCTCTGTGTGGTGGAGAGCTGGCAAGGTCTACTCTTTGACTTC 300
Db 313 CCGGAGGCGAAGAGCGATCTGTGGCGACGGTGAATATCGGCTCCACAAGGGGCTCTGT 372
QY 301 CCGAGGGCGAAGAGCGATCTGTGGCGACGGTGAATATCGGCTCCACAAGGGGCTCTGT 360
Db 373 CTGGATAAGCGGAGCTGGAGACTACATCACTCTCTGAGAGGCGGAGTGAAGGGCTG 432
QY 361 CTGGATAAGCGGAGCTGGAGACTACATCACTCTCTGAGAGGCGGAGTGAAGGGCTG 420
Db 433 CTGGCTGTGGCACCAGCCCGGACCCGACCTGCTGGAACCTGTGTAATGGCACTGTG 492
QY 421 CTGGCTGTGGCACCAGCCCGGACCCGACCTGCTGGAACCTGTGTAATGGCACTGTG 480
Db 493 GTGGCACTTGGCGAGATGAGAGGCTAGCGCCCTTACGCCGAGAGAACTCCCTGTT 552
QY 481 GTGGCACTTGGCGAGATGAGAGGCTAGCGCCCTTACGCCGAGAGAACTCCCTGTT 540
Db 553 CTGTTTGAAGGGGAGAGAGTGTATTCACCATCCGGAAGAGAGATACATGGGAAGATC 612
QY 541 CTGTTTGAAGGGGAGAGAGTGTATTCACCATCCGGAAGAGAGATACATGGGAAGATC 600
Db 613 CCTGTGCTCCGCGCATCCGGGGGAGAGTGAAGTGTACACCACTGATAGTGTATGAG 672
QY 601 CCTGTGCTCCGCGCATCCGGGGGAGAGTGAAGTGTACACCACTGATAGTGTATGAG 660
Db 673 AACCCAGCTCATCAAGCCACCATGTGTACCAAGAGAGAGGCTTAGAGACAGATC 732
QY 661 AACCCAGCTCATCAAGCCACCATGTGTACCAAGAGAGAGGCTTAGAGACAGATC 720
Db 733 TACTACTCTTCCGAGAGAGATCTGTACAAAGAAATCTGAGGCTCTCTCAATGTGTC 792
QY 721 TACTACTCTTCCGAGAGAGATCTGTACAAAGAAATCTGAGGCTCTCTCAATGTGTC 780
Db 793 CGTGTGCGCCAGTGTGAGGGGGAGCCAGGGTGGGAAAGTTCATCTCAAGTCTCCAG 852
QY 781 CGTGTGCGCCAGTGTGAGGGGGAGCCAGGGTGGGAAAGTTCATCTCAAGTCTCCAG 840
Db 853 TGGAAACCTTTTGAAGAGCATGTGTATGAGAGTGTGTGCCACCAACAAGAACTC 912
QY 841 TGGAAACCTTTTGAAGAGCATGTGTATGAGAGTGTGTGCCACCAACAAGAACTC 900
Db 913 AACAGGCTGCAAGAGCTTCTGCTGCTGCTGACCCAGCCAGCGGAGTGGAGAGACAGC 972
QY 901 AACAGGCTGCAAGAGCTTCTGCTGCTGCTGACCCAGCGGAGTGGAGAGACAGCAGG 960
Db 973 GTCTATGCTGTTTCTCCAAACCCCTGAGACTACTAGCCGCTGTGTATTCCTCGGT 1032
QY 961 GTCTATGCTGTTTCTCCAAACCCCTGAGAACTACTAGCCGCTGTGTATTCCTCGGT 1020
Db 1033 GACATTTGACAAGGCTTCCGTAACCTCCACTCAAGGAGCTACCAAGGCTTCCCAAC 1092
QY 1021 GACATTTGACAAGGCTTCCGTAACCTCCACTCAAGGAGCTACCAAGGCTTCCCAAC 1080
Db 1093 CCGGCGCTGCAAGTGCCTCCAGAGCAGAGCGATACCCAGAGACCTTCCAGGTG 1152
|||||

QY 1081 CCGCGGCTGCAAGTGCCTCCAGACAGCCGATACCCAGAGACCTTCCAGGTG 1140
Db 1153 GCTAGCCGTCAACCCAGAGAGTGGCGAGAGGGTGGAGGCCATGGAGGCTTGAAGACCCA 1212
QY 1141 GCTAGCCGTCAACCCAGAGAGTGGCGAGAGGGTGGAGGCCATGGAGGCTTGAAGACCCA 1200
Db 1213 TTGTTCCACTCTTAATACACATACACAGAAAGTGGCGGTTCACCGCATGACAGCCAGCCAC 1272
QY 1201 TTGTTCCACTCTTAATACACATACACAGAAAGTGGCGGTTCACCGCATGACAGCCAC 1260
Db 1273 GGGAGACCTTTCATGTGTCTTACTTAACACAGAGGGGACATATCCACAGTGTG 1332
QY 1261 GGGAGACCTTTCATGTGTCTTACTTAACACAGAGGGGACATATCCAAAGTGTG 1320
Db 1333 GAACGGGGGAGCAGAGAGACAGTGTGGCTTCAACATCATGAGATCCAGCCCTCCGC 1392
QY 1321 GAACGGGGGAGCAGAGAGACAGTGTGGCTTCAACATCATGAGATCCAGCCCTCCGC 1380
Db 1393 CGCGGCTGCTCATCCAGACCATGTGCTGATGCTGAGCGGAGAGAGCTGTATGTAGC 1452
QY 1381 CGCGGCTGCTCATCCAGACCATGTGCTGATGCTGAGCGGAGAGAGTGTATGTAGC 1440
Db 1453 TCCAGTGGAGGTGAGCCAGAGTCCCTGAGCTGTGTAGTCTATGGCGGGGCTGC 1512
QY 1441 TCCAGTGGAGGTGAGCCAGAGTCCCTGAGCTGTGTAGTCTATGGCGGGGCTGC 1500
Db 1513 CAGGTTGGCTCATGTCCCGAGACCCCTACTGCGGGCTGGGAGCAGGGCGCTGCATCC 1572
QY 1501 CAGGTTGGCTCATGTCCCGAGACCCCTACTGCGGGCTGGGAGCAGGGCGCTGCATCC 1560
Db 1573 ATCTACAGCTCCGAAGGTCAGTGTCTCAATCCATTATCCAGCCGAGCCACACAGAG 1632
QY 1561 ATCTACAGCTCCGAAGGTCAGTGTCTCAATCCATTATCCAGCCGAGCCACACAGAG 1620
Db 1633 TGTCCCAACCCCAACACAGAGAGGCCCCACTGCAAGAGTTTCCCTGGCCCCAACTCT 1652
QY 1621 TGTCCCAACCCCAACACAGAGAGGCCCCACTGCAAGAGTTTCCCTGGCCCCAACTCT 1680
Db 1693 CGCTACACCTGAGCTGCCCATGGAAATCCGGCCAGCCCACTACTGAGGCGCACAG 1752
QY 1681 CGCTACACCTGAGCTGCCCATGGAAATCCGGCCAGCCCACTACTGAGGCGCACAG 1740
Db 1753 GAGAAGCTGAGAGAGCTGGCAACTGTGTACAGAGAGCCCAACTGCATCTTTCATC 1812
QY 1741 GAGAAGCTGAGAGAGCTGGCAACTGTGTACAGAGAGCCCAACTGCATCTTTCATC 1800
Db 1813 GAGAAGCTGAGAGAGCTGGCAACTGTGTACAGAGAGCCCAACTGCATCTTTCATC 1872
QY 1801 GAGAAGCTGAGAGAGCTGGCAACTGTGTACAGAGAGCCCAACTGCATCTTTCATC 1860
Db 1873 TTCCGAGAGCTGAGCACTGGCACTGTGCGGAGAGAGGCACTATGAGCGAGACACTC 1932
QY 1861 TTCCGAGAGCTGAGCACTGGCACTGTGCGGAGAGAGGCACTATGAGCGAGACACTC 1920
Db 1933 CTGGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1992
QY 1921 CTGGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1993 CTGGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022
QY 1981 CTGGGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2010
|||||

RESULT 4
LOCUS AB017532 2608 bp mRNA ROD 09-MAR-1999
DEFINITION Mus musculus mRNA for msemk1p, complete cds.
ACCESSION AB017532
NID 94519587
VERSION AB017532.1 GI:4519587
KEYWORDS msemk1p.
SOURCE Mus musculus postnatal brain cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
1 (sites)
Sato, Y. and Takahashi, H.
Molecular cloning and expression of novel Semaphorin family of molecule
Unpublished (1998)
2 (bases 1 to 2608)
Takahashi, H. and Sato, Y.
Direct Submission
Submitted (08-SEP-1999) to the DDBJ/EMBL/GenBank databases. Hiroshi Takahashi, Mitsubishi Kasei Institute of Life Sciences, Neurobiology Lab; 11 Minamiooya, Machida, Tokyo 194, Japan (E-mail: hiroshi@libra.lis.m-kagaku.co.jp, Tel: 81-427-24-6211, Fax: 81-427-24-6314)

FEATURES
source

Location/Qualifiers
1..2608
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="postnatal"
/tissue_type="brain"
5..2000
/gene="msemk1"
6..2000
/gene="msemk1"
/codon_start=1
/product="msemk1p"
/protein_id="BAA75665.1"
/db_xref="PID:d1038405"
/db_xref="GI:4519588"
/translation="MTPPPGRAPASAPRARVLSLPARFLRLRLLEFVWVAASA
OHSRSRGPRIASVMKGODHVDSEPHFVLFHEGSEFVWVGSGKXVHEFPEGN
ASRVYINIGSTSGCODKODGNYITLLERNGNLLGCTNARKSCMNLVNDSTVMS
LCEMGVAPSPDSNLYLFEEDVYSTTRKDEYNGKIPFRFRIGESLITSDTVMQ
NPQFIKATIVHDDAYDDKITYFFREDNEDKNEPAPLNSRYAQLCRGQGGESSLSV
SKWNFLKAMLVSDAATNRNRLQDVFLLDPFSGQWDITVYGVFSNPMYSAVCY
YSLGIDIVERTSLKGYHMG.LPNRPGLCLPKQIPLETFQVADSHPEVNAQREPM
GPLKPLFHSKYHOKVYVYHMQASNGEFLVLTLDGTLHKYVESGDDHSEFVN
IMEIOPFHRAAIOAISLDADRKLVTYSOMESQVPLDMCEVYSGSGCLMSDPMY
GWMDDRCVSTISORSVYOSTINPAEPHECNPBPDEAPLQKVSILARSKRYLLTCM
ESRHAITYLMRHEENVQSCPEPHOSPSCILFIENITARYGHTKCEADQGSTLREAOH
WELLPEDRALAEOLMGMHARALASFWLGVLPFLILGLLVH"

BASE COUNT 554 a 769 c 727 g 558 t
ORIGIN

Query Match 71.8% Score 1444: DB 32: Length 2608:
Best Local Similarity 87.4% Pred. No. 0.00e+00: Mismatches 241: Indels 12: Gaps 2:
Matches 1757: Conservative 0:

Db 6 ATGACGCTCTCTCTCCGAGAGTGCAGCCGCCAGCGACACCGCGCCGCTCTCAGC 65
OY 1 ATGACGCTCTCTCTCCGAGAGTGCAGCCGCCAGCGACACCGCGCCGCTCTCAGC 60
Db 66 CTGCGCGCTCGTTCGGGCTCCCGCTGCGGCTCGGCTTCTGCTGTTCGGGTGCGC 125
OY 61 CGCGCGCTCGTTCGGGCTCCCGCTGCGGCTCGGCTTCTGCTGTTCGGGTGCGC 120
Db 126 GCGGCTCGGCGCAAGGCACTCGAGAGCGGACCCCGCATCTCCGCGCTGGAAGG- 184
OY 121 GCGGCTCGGCGCAAGGCACTTAAGAGCGGACCCCGCATCTTCGCTTGGAAAGGC 180
Db 185 -----GCAGGACCATGTGACTTATGACACCTGAGCCACACACCGTGTTCAT 236
OY 181 CATGTAGGCGAGGACCGGCTGACTTTCGCGAGACTGAGCCGACACCGGTGCTTTCCAC 240
Db 237 GAGCGGGGAGCTTCTCTGTTCGGGTGAGCGTGGCAAGGTATACCATCTCACTTC 296
OY 241 GAGCGGGGAGCTTCTCTGTTCGGGTGAGCGTGGCAAGGTATACCATCTCTTTCGCTTC 300
Db 297 CCGGAGGCGAAGATCCCTCTGTGCGCAGGAGTGAACATGGGCTCCCAAGAGGGTCTCT 356
OY 301 CCGGAGGCGAAGATCCCTCTGTGCGCAGGAGTGAACATGGGCTCCCAAGAGGGTCTCT 360

Db 357 CAGGACAAACAGGACTGTGGAATTACATCAGCTTCTGAAAGCGGGGCTAATGGGCTG 416
OY 361 CTGGATTAAGCGGACTTCGAGAACTATCACTCTCTCTGAGAGGGGAGGTAGGGGCTG 420
Db 417 CTGGTCTGTGGCAACAATCCCGGAAGCCAGCTGTGAGACTTGTGATGACAGTGTG 476
OY 421 CTGGCTGTGGCAACAAGCCCGGACCCAGCTGTGAGACTTGTGATGAGCTGTG 480
Db 477 GTATGTCTACCTGTGTAGATGAAGGCTATGCCCCCTTAGCCCCGATGAGACTTCCTG 536
OY 481 GTGC---CACTTGTGGCAATGAGAGGCTACGCCCTTACGCCGAGCAAGACTCCCTG 537
Db 537 GTTCCTTTTGAAGGAGATGAAGTCTCTTACATCCGGAACGAGATATCAACGGAAG 596
OY 538 GTTCTTTTGAAGGAGGACAGAGTGTATTCACCATCCGGAACGAGATATCAATGGAG 597
Db 597 ATCCCTCGGTTTGCAGCATTCGGGGCGAGTGAAGTGTACACAAGTGTATCAGTCATG 656
OY 598 ATCCCTCGGTTTCCGGCGCATCCGGGGCGAGTGTATGAGCTGTACACAGTATCTGTGATG 657
Db 657 CAGAACCCACAGTTCTATCAGGCCACCATTTGTACCAAGACCAAGCCTATGATGATAAG 716
OY 658 CAGAACCCACAGTTCTATCAAGCCACCATCTGTGCACCAAGACAGGCTTACGATGACAAG 717
Db 717 ATCTACTTCTTCCGAGAGACAAACCTGACAAAGAACCCCGAGGCTCTCTCAATGTG 776
OY 718 ATCTACTTCTTTCGAGAGGACAACTCTGACAAAGATCTGAGAGCTCTCTCAATGTG 777
Db 777 TCCCGATGAGCCAGTTGTGAGGGGGGACAGGGTGTGAGAGTTGTTGTTCTCTCC 836
OY 778 TCCCGATGAGCCAGTTGTGAGGGGGGACAGGGTGTGAGAGTTGTTGTTCTCTCC 837
Db 837 AAGTGAACACTTCTCTGAAGCCATGTTGTTCTGACAGGATGACCAACCAAGAAC 896
OY 838 AAGTGAACACTTCTCTGAAGCCATGTTGTTCTGACAGGATGACCAACCAAGAAC 897
Db 897 TTCAATCGGCTCAAGATGTTCTTCTGCTCCCTGACCCAGAGGGGAGGAGATATCC 956
OY 898 TTCAACAGGCTCAAGAGCTCTCTCTGCTCTGACCCAGGCGCAATGAGGAGACAC 957
Db 957 AAGGCTATGAGGCTTCTTCCAAACCCCTGGAACCTACAGCTGTGCTGTATTCCTT 1016
OY 958 AAGGCTATGAGGCTTCTTCCAAACCCCTGGAACCTACAGCTGTGCTGTATTCCTT 1017
Db 1017 GGTGACATTGAGAGTCTTCTGCTACCTCATGCTTCAAGAGGTACCAATGGGCTTCC 1076
OY 1018 GGTGACATTGAGAGTCTTCTGCTACCTCATGCTTCAAGAGGTACCAATGGGCTTCC 1077
Db 1077 AACCTCGACCTGGCATGTGCTTCCCAAAAAAGAGCCATACCCAGAAACCTTCCAG 1136
OY 1078 AACCCGGGCTCGGCAAGTGTCTCCCAAGACAGAGCCGATACCCAGAGAGCTTCCAG 1137
Db 1137 GTAGCTGATAGTACACAGAGGTGCTCAGAGGGTGAAGACTATGGGCCACTGAAGCA 1196
OY 1138 GTAGCTGATAGTACACAGAGGTGCTCAGAGGGTGAAGACTATGGGCCACTGAAGCA 1197
Db 1197 CCATTGTTCCATTCAATGATACCATTAACCAAGAAAGTGTCTTACCGCATGCAAGCAGC 1256
OY 1198 CCATTGTTCCATTCAATGATACCATTAACCAAGAAAGTGTCTTACCGCATGCAAGCAGC 1257
Db 1257 AATGAGAGACCTTCATGCTGTTATATCAACACAGAGAGGGGACATTCACAAAGGTG 1316
OY 1258 CAGGAGGAGACCTTCATGCTGTTATATCAACACAGAGAGGGGACATTCACAAAGGTG 1317
Db 1317 GTGAAATCAGGAGGACAGACCATAGCTTGTCTTCAATATATCATGAGATCAGGCTT 1376
OY 1318 GTGAAATCAGGAGGAGACAGACCATAGCTTGTCTTCAATATATCATGAGATCAGGCTT 1377
Db 1377 CACGCTGAGAGGCGCATTCAGGCTATATCATTTGATGCTGACCGGAGAACTATATG 1436
OY 1378 CCGCGCGGCGGCTGCGCATCAACATGCTGTGATGCTGAGCGAGGAACTGTATGTG 1437
Db 1437 ACCTCCAGTGGGAAGTGAAGCAGGTACCTCGACATGTGTGAGGCTTACAGGCGGGCG 1496

| | | | | |
|------------|--|--|-------------------------|-----------------|
| Oy | 1438 | AGCTCCAGTGGGAGGTAGGCCAGGTGCCCTCTGAGAC | TGTGTGAGTGTATGTGCGGGGGC | 1497 |
| | | | | |
| Db | 1497 | TGCCATGGCTGCCTCATGTCCCGAGACCCCTTACTGTGCTGGGACCGAGCCGTGCGTG | | 1556 |
| Oy | 1498 | TGCCACGGTGTGCTCATGTCCCGAGACCCCTTACTGTGCTGGGACCGAGCCGTGCGTG | | 1557 |
| | | | | |
| Db | 1557 | TCTATCTACAGTGTCCCAAGGATCGTGTGCGAGTGCATTTAATCCAGGGAGCCACACAGA | | 1616 |
| Oy | 1558 | TCCATCTACAGTGTCCCAAGGATCGTGTGCGAGTGCATTTAATCCAGGGAGCCACACAG | | 1617 |
| Db | 1617 | GAGTGTCCCAACCTTAACACAGACGAGCGCTCATTCAGAAAGTTTCCCTGGCTCGAAC | | 1676 |
| Oy | 1618 | GAGTGTCCCAACCCCAACACAGACGAGCGCTCATTCAGAAAGTTTCCCTGGCCCAAC | | 1677 |
| | | | | |
| Db | 1677 | TCTGATACCTACCTGACCTGCCCCATGGAGTCCCGCCACGCGACCTTACTTATGCGCCAT | | 1736 |
| Oy | 1678 | TCTGCTACTACCTGACCTGAGCTGCCCATGTGAAATCCCGCCACGCGACCTACTCATGCGCCAC | | 1737 |
| | | | | |
| Db | 1737 | GAGGAGATGTGTGAACACAGAGTGTAGCCAGGCGCACCAAGCCCTTACGTGCATCCTGTTC | | 1796 |
| Oy | 1738 | AAGGAGAACGTGTGACAGAGCTGCGCAACCTGGTACACAGAGCCCAACCTGATCCTGTTC | | 1797 |
| | | | | |
| Db | 1797 | ATTGAGACCTACAGCGCCCGCTAGTATGTAGCCACTACCGCTGGAGAGCCCAAGAGGCTTC | | 1856 |
| Oy | 1798 | ATCGAGAACCTACAGCGCCCGAGAGTATGCGCCACTACTTCTGGGAGCCCGAGAGGGCTCC | | 1857 |
| | | | | |
| Db | 1857 | TACCTCGGTGAGGCTCAACACTGTGGAGCTGTGCCAGAGACAGACACTGGCGCAACAA | | 1916 |
| Oy | 1858 | TACTTCGCGAGAGCTACAGACTGTGAGAGCTGTGCCAGAGACGGCATATGCGCCAGCAC | | 1917 |
| | | | | |
| Db | 1917 | CTAATGGCCATATGCCCGGGCCCTGTGGCGCTCTCTGTGCTGTGGGGTCTCTGCCACACTC | | 1976 |
| Oy | 1918 | CTGCGGGTCAAGCTCGTGTGCCCTGTGCCCGCTCTCTGTGCTGTGGGGTGTGTGCCACACTC | | 1977 |
| | | | | |
| Db | 1977 | ATACTTGGTCTGCTGGTCACTACAGGGGCTTC | | 2006 |
| Oy | 1978 | ACTTGTGCTGTGCTGTCCACTAGGGGCTTC | | 2007 |
| | | | | |
| RESULT | 5 | | | |
| LOCUS | AF030699 | 1192 bp | mRNA | ROD 04-SEP-1998 |
| DEFINITION | Mus musculus semaphorin L (Semal) mRNA, partial cds. | | | |
| ACCESSION | AF030699 | | | |
| NID | 93523116 | | | |
| VERSION | AF030699.1 | GI:3523116 | | |
| KEYWORDS | | | | |
| SOURCE | house mouse. | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| AUTHORS | 1 (bases 1 to 1192)
Lange, C., Liehr, T., Goen, M., Gebhart, E., Fleckenstein, B. and Enser, A. | | | |
| TITLE | New eukaryotic semaphorins with close homology to semaphorins of DNA viruses | | | |
| JOURNAL | Genomics 51 (3), 340-350 (1998) | | | |
| MEDLINE | 98389619 | | | |
| REFERENCE | 2 (bases 1 to 1192)
Enser, A., Liehr, T., Lange, C., Goen, M., Gebhart, E. and Fleckenstein, B. | | | |
| AUTHORS | Direct Submission
Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare Virologie, Friedrich Alexander Universitaet, Schlossgarten 4, Erlangen, Bavaria 91054, Germany | | | |
| FEATURES | Location/Qualifiers | | | |
| Source | 1..1192
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="9A3..3-B"
/chromosome="9"
1..>1192
/gene="Semal" | | | |
| gene | | | | |

| | | | | |
|-----------------------|--------------|---|-----------------|--------------------|
| CDS | | /evidence=not_experimental
13.. >1192
/gene="Semal"
/note-"W-Sema-L"
/codon_start=-1
/evidence=not_experimental
/product="semaphorin L"
/protein_id="AAC34262.1"
/db_xref="PID:g3523117"
/db_xref="GI:3523117" | | |
| BASE COUNT | 254 a | 355 c | 336 g | 247 t |
| ORIGIN | | | | |
| Query Match | 42.1%; | Score 847; | DB 32; | Length 1192; |
| Best Local Similarity | 88.1%; | Pred. No. 0.0e+00; | | |
| Matches 1048; | Conservative | 0; | Mismatches 129; | Indels 12; Gaps 2. |
| Db | 13 | ATGAGCGCTCCTCCCTCCGGACGTGCCGGCCCCCGACGGCACGGCGCGCGCGGCTTGCAGC | 72 | |
| Oy | 1 | ATGAGCGCTCCTCCGGACGTGCCGGCCCCCGACGGCACGGCGCGCGCGGCTTGCAGC | 60 | |
| Db | 73 | CTGCGCGCTCGGTTTCGGGACTCCCCTCGGCTCGGCTTGCTGCTGTGTGGTGGCC | 132 | |
| Oy | 61 | CCGCGCGCTCGGTTGGGCTTCCGCTCGGCTCGGCTTGCTGCTGTGTGGTGGCC | 120 | |
| Db | 133 | GCCGCTCCGCCCAAGGCGCACTGAGAGCGGACCCCGCATCTCCGCGTCTGGAAGG- | 191 | |
| Oy | 121 | GCCGCTCCGCCCAAGGCGCACTGAGAGCGGACCCCGCATCTCCGCGTCTGGAAGG | 180 | |
| Db | 192 | -----GCAGAACATGTGGATTATTAAGCAGCTTAGCCACACACCCTGCTTTCCAT | 243 | |
| Oy | 181 | CATGTAGGCGAGGACGGGTGGATTGGCCAGACTAGCGCGCACAGGCTGCTTTCCAC | 240 | |
| Db | 244 | GAGCGGGGACACTTCTCTGTGTGGGTGGGTGGAAGTGGTCAACCTTCACTTC | 303 | |
| Oy | 241 | GAGCGGGGACACTTCTCTGTGTGGGTGGGTGGAAGTGGTCAACCTTCACTTC | 300 | |
| Db | 304 | CCCGAGGGCAAGAAATGCCCTCTGTGTGCGACGGTGAACAATCGGCTCCACAAAAGGGTCTGT | 363 | |
| Oy | 301 | CCCGAGGGCAAGAAATGCCCTCTGTGTGCGACGGTGAACAATCGGCTCCACAAAAGGGTCTGT | 360 | |
| Db | 364 | CAGGACAAACAGACTGTGGGAATTACACTCTTTAGAAAGCGGGGTAATGGCTG | 423 | |
| Oy | 361 | CTGGTAAAGCGGACTGTGGGAATTAATCACTCTCTGTGGAGAAGCGGAGTGAAGGGCTG | 420 | |
| Db | 424 | CTGGCTGTGGCAACATGCCCGGAACCCAGCGTGGGAATGGGAATGACAGTGTG | 483 | |
| Oy | 421 | CTGGCTGTGGCAACATGCCCGGAACCCAGCGTGGGAATGGGAATGACAGTGTG | 480 | |
| Db | 484 | GTGATGTACTTGTGTGATGAAGAGGTATCCCTCTTACGCCCGGATGAGAACTCCTG | 543 | |
| Oy | 481 | GTGC---CACTTGGCGAATGAGAGGCTACGCCCTTTCAGCCCGGAGAGAACTCCTG | 537 | |
| Db | 544 | GTTCTGTTGAAGAGATGAGTGTACTTACCATCCGGAAGCAGGAATACACGGGAG | 603 | |
| Oy | 538 | GTTCTGTTGAAGAGAGATGAGTGTACTTACCATCCGGAAGCAGGAATACATGGGAG | 597 | |
| Db | 604 | ATCCCTGSGTTTCGACGATTCGGGGGGAAGTGAATGTAAACAATGATTAACGTATG | 663 | |
| Oy | 598 | ATCCCTGSGTTTCGACGATTCGGGGGGAAGTGAATGTAAACAATGATTAACGTATG | 657 | |
| Db | 664 | CAGAACCACAGTTCATCAAGAGCCACCATTTGTGCACCAACCAAGGCTATGATGAAG | 723 | |
| Oy | 658 | CAGAACCACAGTTCATCAAGAGCCACCATTTGTGCACCAACCAAGGCTATGATGAAG | 717 | |

```
Db 724 ATCTACTACTTCTTCGAGAGAACACACCTGACAGAACCCCGAGCTCTCTCATGTG 783
|||||
Qy 718 ATCTACTACTTCTTCGAGAGAGACATCTGACAGAAATCTGTAGGCTCTCTTAATGTG 777
|||||
Db 784 TCCCGATAGCCCGAGTGTGTGAGGGGAGACAGGCTGTGTAGATGTGTTGTGTCTCC 843
|||||
Qy 778 TCCCGTGTGGCCGAGTGTGTGAGGGGAGACAGGCTGTGTAGATGTGTTGTGTCTCC 837
|||||
Db 844 AAGTGAACACCTTCTCGAAAGCCATGTGTGTGACGAGCGATGACACCAAGAGAAC 903
|||||
Qy 838 AAGTGAACACCTTCTCGAAAGCCATGTGTGTGACGAGCGATGACACCAAGAGAAC 897
|||||
Db 904 TTCAATGCGTGCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 963
|||||
Qy 898 TTCAACAGGCTGCAGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 957
|||||
Db 964 AGGCTCATAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023
|||||
Qy 958 AGGCTCATAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
|||||
Db 1024 GTGACATTTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1083
|||||
Qy 1018 GTGACATTTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1077
|||||
Db 1084 AACCTTGACCTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1143
|||||
Qy 1078 AACCTTGACCTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1137
|||||
Db 1144 GTGACATTTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1192
|||||
Qy 1138 GTGACATTTGACAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1186
|||||

RESULT 6 AF030697 8861 bp DNA PRI 09-SEP-1998
DEFINITION Homo sapiens semaphorin L (SEMA) gene, partial cds.
ACCESSION AF030697
NID 93551778
VERSION AF030697.1 GI:3551778
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Lange C., Liehr T., Goen M., Gebhart E., Fleckenstein B. and
Ensser A.
New eukaryotic semaphorins with close homology to semaphorins of
DNA viruses
JOURNAL Genomics 51 (3), 340-350 (1998)
MEDLINE 98389619
REFERENCE
AUTHORS 2 (bases 1 to 8861)
Ensser A., Liehr T., Lange C., Goen M., Gebhart E. and
Fleckenstein B.
Direct Submision
Submitted (22-OCT-1997) Institut fuer Klinische und Molekulare
Virologie, Friedrich Alexander Universitaet, Schlossgarten 4,
Erlangen, Bavaria 91054, Germany
FEATURES
Source
1..8861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="15p22.3-p23"
/chromosome="15"
join(<1..114,605..646,944..1036,1238..1322,1443..1553,
2176..2315,2907..3091,3945..4053,4146..4344,6876..7003,
7178..7332,7533..7594,7904..>8265)
/gene="SEMA"
/product="semaphorin L"
/evidence="experimental"
<1..>8265
/gene="SEMA"
join(<1..114,605..646,944..1036,1238..1322,1443..1553,
CDS
```

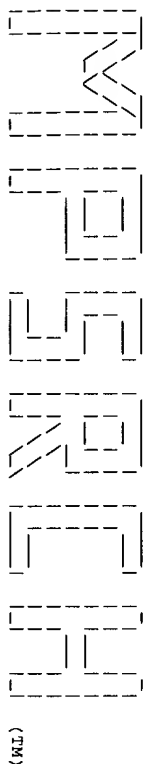
```
2176..2315,2907..3091,3945..4053,4146..4344,6876..7003,
7178..7332,7533..7594,7904..8265)
/gene="SEMA"
/sequence starts with the putative second exon of
SEMA: H-SEMA-L"
/codon_start=1
/product="semaphorin L"
/protein_id="AAC34741.1"
/db_xref="PIR:93551779"
/db_xref="GI:3551779"
/translation="EPHTVLEFEPSSWVYGGKGYLLPDPPEKNAVPTVNIQST
KSCDRLDCEVYITLLERBSGLACTTNARHPSCHLVNLTGYPLCEMGIAPFS
DENSTVLEGGDVTSTIKOEINQITPFRIRIGSELITSDTVNQPFKATIVQ
DOAYDKITVFFREDNPDKNPEAPLVNRYVQLOCGDGGSSLSVSKMFLKMLV
CSDAATNKNFNRLQVFLPDPGQMRVRYGVSNPNVNSAVCVSLGIDIDKVFRT
SLKGYHSLSNPBPQCLPPOOPIPTFTFOVADHPVQARVEPMGLKPLFHSKY
HYOKVAVHMQASHEGTFFHLVLTIDRGTIRKVVPEGDSHFAFNIMETOPFRRAA
IOTMSLDAERKLYSSOMEVSOVPLDCEVYGGGCHCLMSRDPYCGWDGRCISTY
SSERSVLOSINPAEPHKECPNRPDKAPLVQVSLAPNSRYLLSCPMESRHATYSMRK
ENVEDSCFPGHQSPTCLIFENLTAQOYGHVCEAEGSYFREAOHMLPDEIDMAE
HLGHACALAAASLWLVGYLPTLTLGLVH"
BASE COUNT 1749 a 2358 c 2757 g 1997 t
ORIGIN
Query Match 18 7% Score 375; DB 31; Length 8861;
Best Local Similarity 100.0%; Pred. No. 9,69e-270;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 7900 CCAGACAAAGGCCCACTGAGAGGTTCCCTGGGCCCAAACTCGCTACTACTGTAGC 7959
|||||
Qy 1636 CCAGACAAAGGCCCACTGAGAGGTTCCCTGGGCCCAAACTCGCTACTACTGTAGC 1695
|||||
Db 7960 TCCCCATGGAATCCCGCCAGCCACCTACTATGAGCCGACAGAGAGAGCTGAGCAG 8019
|||||
Qy 1696 TCCCCATGGAATCCCGCCAGCCACCTACTATGAGCCGACAGAGAGAGCTGAGCAG 1755
|||||
Db 8020 AGCTGGGAACCGGTGACACAGAGCCCACTACTGCTGTTTCATGAGAACCTCAGCGCG 8079
|||||
Qy 1756 AGCTGGGAACCGGTGACACAGAGCCCACTACTGCTGTTTCATGAGAACCTCAGCGCG 1815
|||||
Db 8080 CAGCAGTACGGCCACTACTTCTGCGAGGCCGAGAGGCTCCTACTTCGGAGGCTCAG 8139
|||||
Qy 1816 CAGCAGTACGGCCACTACTTCTGCGAGGCCGAGAGGCTCCTACTTCGGAGGCTCAG 1875
|||||
Db 8140 CACTGGCAGCTGCTGCCGAGAGCGCATCATGCGCGAGCACTGCTGCTATGCTCTGT 8199
|||||
Qy 1876 CACTGGCAGCTGCTGCCGAGAGCGCATCATGCGCGAGCACTGCTGCTATGCTCTGT 1935
|||||
Db 8200 GCCCTGGCGGCTCTGCTGTGGGGGTGCTGCCACACTCCTTGGCTTCTGTGTC 8259
|||||
Qy 1936 GCCCTGGCGGCTCTGCTGTGGGGGTGCTGCCACACTCCTTGGCTTCTGTGTC 1995
|||||
Db 8260 CACTAGGCGCTCCCG 8274
|||||
Qy 1996 CACTAGGCGCTCCCG 2010
|||||

RESULT 7
LOCUS AHU18243 11874 bp DNA VRL 30-SEP-1995
DEFINITION Alcelaphine herpesvirus 1 putative semaphorin homolog (Ahv-sema)
and putative membrane antigen genes, complete cds, and major
ssDNA-binding protein gene, partial cds.
ACCESSION U18243
NID g1000716
VERSION U18243.1 GI:1000716
KEYWORDS
SOURCE
ORGANISM Alcelaphine herpesvirus 1.
Alcelaphine herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae; unclassified Gammaherpesvirinae.
REFERENCE
AUTHORS Ensser A. and Fleckenstein B.
TITLE Alcelaphine herpes type 1 has a semaphorin-like gene
```


JOURNAL J Gen. Virol. 76 (Pt 4), 1063-1067 (1995)
 MEDLINE 97201573
 REFERENCE 2 (bases 1 to 11874)
 AUTHORS Ensser, A.
 TITLE Direct Submission
 JOURNAL Submitted (07-DEC-1994) Armin Ensser, Institut fuer Virologie,
 Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten
 4, Nuernberg, Bavaria 91054, Germany
 FEATURES
 source
 1..11874
 /organism="Alcelaphine herpesvirus 1"
 /strain="C500"
 /specific_host="Carnochaeetes taurinus taurinus"
 /db_xref="taxon:35252"
 /lab_host="Bos taurus"
 /clone="clones C38, C249, C500Bs"
 <1..(330,346)
 /note="gamma-2 herpesviral GC-rich terminal tandem
 repeats"
 /function="putative packaging signals"
 /rpt_type=tandem
 3573..5534
 /standard_name="AHV-sema"
 /note="putative semaphorin homolog; similar to Genbank
 Accession Number L26081"
 /codon_start=1
 /protein_id="AAC54475.1"
 /db_xref="PID:g1000717"
 /db_xref="GI:1000717"
 /translation="MAYLNATKSPVLSILSKVYLFKFGGEGGCGAGLITFEVYH
 PAMGTICVSTRILMLTISATTAASKRPIDKPRILVITDFGCGHFRGPEPRHTVLFH
 SLNSDDYVGNNNTYLPDFAHSSMASTALINTSHNTIRLSTGCNFTYLLHNTD
 GLAACISQKPSQCMILNNLTQFLFGKALAEFSSGMLVDFDNDITSLTKS
 LSSGKRRRIAGQVEIYLTSDTAMHRLPGQATVAHKESVDDEIYDFDONSDFKO
 EPHTVRGVGVCSDDGSESLSVYKWTFLKARLACVDVDTGRIYELDIFIMWAP
 ENMEELIYGLFLSPWNPFSACVCFYKTDIDHYKTSKLYNHHKLTPTPRGCMKH
 QHPTEFYADRIPEVADPYVQKNNAPFIISKYITLYTLVRYEGVFNATIFY
 LTTIKGTHLYVRDEDSNTTALNLTINLTFNFOKPAQIQLLNDTNLTKLYNEMES
 EYPLDLCVYGNDCSCFMSRDPACTYNNTKSCQSVETGEPAPRTLSEMGDHY
 APYVVKHQSILPLLSNGYLSCPAVSNADYFKDGTGERCHVKTKNDICILLNS
 TPAENGHCNMEDSVYKLFENVYTLM"
 6414..10523
 /note="putative membrane anti-gen: similar to p140 (PIR
 Accession Number S29606), and p160 (Genbank Accession
 Number Z26584)"
 /codon_start=1
 /protein_id="AAC54476.1"
 /db_xref="PID:g1000718"
 /db_xref="GI:1000718"
 /translation="MTRLEQLLVAYTAPEPSALETTAALRSIPGLONTYIOTDAYL
 VFEFTSRPRREHOLKIELFLIKALRTDEDOYLPISALSSSTFEYVGDINL
 PTQSNELPAILQTRDAKEKVFENVERIQRLLFEFGSGSEOLNHTYALLRELLGD
 LTYLQPRNELLFNSLDEYVATYIPRELRKVRHYVGINLNTGASVYNNASPT
 AVYDSEIRPIPTHYVREIDRMETRAVVPVYGHSGVYFRHGGAAVYCSLNTST
 GAGCGDKGMTIFSTFISQPLTGQGLITPEPCGWQDELSTALQMLTKIANTVEARFK
 TLVAGVPLVQGVFTLSPTLREKRAMAFSTSLITGFLPFLAPTAETLDAMERNR
 QLLLEVGEYDYGKGLAHNVNINSSGSHALHLPALQQLMIIPFGVYIICDM
 QYNEKVTMFKIARACGLDGSMSARSIPARTRLKKMVRNNYNNRIITKSNLKD
 SAAALVYIADNEVDYSKMMDIALSGCGPHILGNTPSNTIYISDKNGQYEIYDI
 QYKNNHKKQSEGETEDPPLAPDMSNIDLFKNLVDTEDLLQVLRHPYVGCRAKIHVY
 DRCNNHIAQOPGVGPDPILPCLDSVYVNVLVGDIREGSESVPRVAAMWARRARILVY
 DTFIIFQNTQLSVSPVSRLLKHNWTLSSKVALYLVAEDVYLEBETSYGN EHVLDV
 PYSLTQSTIRPNVLVALLPQCGPDLALALNSGSDPTVYVPGKYRHNREDAI
 PGSTADADFIAGIYVGLSSNIDSVGDSITROMLVNVRQYINDVRRMLKAGSTLAI
 GQALCRILFATKAIQFPGASQOFTPLLPNSRRESSEWLNFKIIPEDIKAAVFDLRC
 VLQWQGHLLGFSNNITFEGLTFETGTOOVAENGLVSCGAREVLPALPTAEHY
 AGLCSEGGRLHALLPCLCALFNTQWQHNTAVTGQOGLFVPSCKMFLRYLYNSKFKQ
 HYSNLTLNRIHTFNEFTQQLDPRHAKRAVADPNQYIIPKAMDPQO"

RESULT 8 AF005370 130608 bp DNA VRL 21-AUG-1997
LOCUS Alcelaphine herpesvirus 1 L-DNA, complete sequence.
DEFINITION AF005370
ACCESSION g2337967
MID AF005370.1 GI:2337967
VERSION
KEYWORDS
SOURCE . wildebeest herpesvirus.
ORGANISM Alcelaphine herpesvirus 1
vires: dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae.
REFERENCE 1 (bases 1 to 130608)
AUTHORS Ensser, A., Pfanz, R. and Fleckenstein, B.
TITLE Primary structure of the alcelaphine herpesvirus 1 genome
J. Virol. 71 (9), 6517-6525 (1997)
MEDLINE 97404659
REFERENCE 2 (bases 1 to 130608)
AUTHORS Ensser, A., Pfanz, R. and Fleckenstein, B.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1997) Institut fuer Klinische und Molekulare
Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg,
Schlossgarten 4, Erlangen 91054, Germany
FEATURES
source
1. 130608
/organism="Alcelaphine herpesvirus 1"
/strain="C500"
/specific_host="Connochaetes taurinus taurinus"
/db_xref="taxon:35252"
/lab_host="MDCK bovine epithelial kidney cells; ATCC
CCL-22"
918. 1211
/note="putative ORF"
/codon_start=1
/evidence=not_experimental
/product="A1"
/protein_id="AAC58052.1"
/db_xref="PID:g2337968"
/db_xref="GI:2337968"
/translation="MRLHFLYRSIEFMYKRSRHKIFCIITSLVFLTAHKRRPARNP
IDPFSRWGCPFCYKRSKQHLIVNOLMSPKNCSLTSAQCITANEDTTK"
1248. 1254
/evidence=not_experimental
complement(1300..1305)
/note="spliced cDNA detected by 5'RACE-PCR"
/codon_start=1
/evidence=experimental
/product="A2"
/protein_id="AAC58053.1"
/db_xref="PID:g2337969"
/db_xref="GI:2337969"
/translation="MSQNSSENPSRKRKYVKMCDLTFEOKERRRSINRASKNFLEK
RRRIPEOQKGLINIKYENSRKRCQVEKKEIIRLEWLVHKCTLLQNTGPPE
PKVKENSLEMOCAFAFLNLDQQTNNININLETYSGNNTINGFAAATATLTNTCEK
TLANNNTNFEAKLNCLEVLPSFTSALDDLISIDMNNLYN."
3492. 5453
/note="A3: AHV-sema, similar to Vaccinia A39"
/codon_start=1
/evidence=not_experimental
/product="semaphorin homolog"
/protein_id="AAC58054.1"
/db_xref="PID:g2337970"
/db_xref="GI:2337970"
/translation="MAYLNATYSKPVISLSISKVYLFENCGGSGCIGLITTEVIVH
PAMGGLVSIIRLILWLSITAKSRFIDKPRILVNLTDGFGQAFEPQEPHEVTLFH
SLNSTDVYVGNNITLYLDFEHAHNSASNAALINITSHTNHRSSCEPITLHNTQD
GLACGTINQSKSCWLINMLITQFLGPKLAPFSPSSGSLVLPQNDITYSTINLYKS
LSGSHKFRITAGOVELYTSDTAMHRPOFOAFAVKNSGYDKIYEFPOENSHSPFO
EPHTYPRVGQVCSDDQGESLSVYKMTTELKARLACVDYDGRIRYNELODFTIQAP
ENSMEETLITGIFLSPMNFSAVCVTVVADIDRVLTSKLNHKLPTPRGQCAKNH
OHVPTETFOVADRIPEVADPVYQKNNAPEIIQSKRYITKLLVYVEVGGVMAQKIFY
LTGTITIHIVREDNSSTALNITLLEINPFOKAPAPIONILIDNTNLKLYNSEWEVS

polyA_signal
CDS
EVPDLICSVYNDGCFSCFMSRDLCTWYNNTCSEKORYSEVETGPGANRTLSEMGDHY
APTVYKHQVSIPLNSYSLSCPAVSNHADYFPTKGFTEKROHVATHKNDICLLGANS
TPTATNGTHYCNMKEDSVYKLLLEVNTLM"
5590. 5597
5732. 6097
/note="ORF containing putative signal peptide: detected by
5'RACE-PCR"
/codon_start=1
/evidence=experimental
/product="A4"
/protein_id="AAC58055.1"
/db_xref="PID:g2337971"
/db_xref="GI:2337971"
/translation="WVAOLYHHLITLISIVYIFPVNAPRTLEDEDEDNSTEVDF
LDSDFQPPRRMSNCNCREZATRTLLRLKRALQIINGINMETDEUPTTPPTMTT
PLRTPLDTSPPVLPSPALP"
6123. 6128
/evidence=not_experimental
6333. 10442
/note="ORF03; similar to H. salmieri and EHV2 ORF3, similar
to ORF75"
/codon_start=1
/evidence=not_experimental
/product="tegument protein"
/protein_id="AAC58056.1"
/db_xref="PID:g2337972"
/db_xref="GI:2337972"
/translation="WITRELOLAYTAEPALSETNAIALNSIPQLQNVIIOTDAYL
VTFSPRPLRHQKLETLFIKALTFPEOQYLPPLASLSSFTFVGPDIRL
PTQGNELTALIQTRDAREKVENYRIHCHLLEFSGSPSQLNHTHALREILCGD
LTLQYFPRNELFSSLDFAVATYPPRFKGLVRYHVGIDINLITGASVNYMSPHT
AVTDSEIFRIPRTYHVRFRIDREIRHVRDVGFGHSRCGRFROGGAAYMSTNST
GATGCLQKGMTSPISOPTLQGLITTPKMGWGOEISATAIQMLTRYANTVEFAK
TLNTAGVPLVGSEVTLSPTLREKRMATGSLITGLPPIYLLKPTAETTLDAEMNR
OLLIVGYPDYGGKGLHPVINNSSEGRHAHLIQAOLRLMIIIPRGKYLHCDDW
QYNEKVTMFKIAKAGDGLSMSARSLPAHRLRLKKNFNLYNRIIKNSMLKVD
SAAALVVIADNEYODVSKMDIALSGMCCPHILGNLTPNSNTIYISDKNOYGETVDI
OYKMHMPKQSPGTEPPLADPMNSIOLEFKLMDTEDLLLOVLRHPYCKGAHIVHY
DRGNGHIAQOPGVGPPDIPCLDSFVYTHNVDDDIRGMSYRVMAADRYARLLI
ETQYSTPELDITDALNNLGGTLTHIREHGVATYSKRYGMCIGTEGTITQDDPL
GTLIVSGCTNCILIGEVNTEEPFLIGISVPEGIYHRAHENSIMAAKDFCSMNF
GFQVNSANGCGLNSVQALNADVPVGPGLKPYFKPGSAIIRVNLHTPEFSGGI
CCMASGASGETFPTPSVALNLOFMLIVAENMLALSHGDSGLICAVCEMFFAG
GLSARLLIHDEDEBPPLFSETGPFGEVVAIDVAIYAALIAIRANLYNEVCIOIGEVVS
DFTVYHONTOLLSVPSYRIRKHNHTJFSKSDLVLYKEDQVLPBETSGNRYEHLTD
PYSTISOSTTPNNVLAHLPPCGGTPDALALITNSGSPDVTIVPGCKYRINNREDA
GSPPTADDFIAGIVLYGSSNIDSDVGDSTIRQWLVNVRQVINDVARNLKAGSFTLAI
GOLACRILIFATKAIGFDAGSOOTPELLPNASRRYSRLNFKIPEDTKAVAFRLRC
VLPCWVGTHLGFSSHNNTTFPGDLETRQOVATNGPLVQSGPAREVPLNPTAEHPY
AGLCEGDSRHLALLEPDLAFNTQWOMONOTGPGQELPVSPMKLMFYRLNMSKPHQ
HYRSLRNLNHTNFEHQQDLDFHADRGAVPHDNPQYIPMDAMDSQ"
11134. 11139
/evidence=not_experimental
11260. 14643
/note="ORF06"
/codon_start=1
/evidence=not_experimental
/product="major ss DNA binding protein"
/protein_id="AAC58057.1"
/db_xref="PID:g2337973"
/db_xref="GI:2337973"
/translation="MALKHOLNHAVEDNLGSKAPIGCGFIYIPETHNFKEISLIG
DKYAEAGFSPILHAGTVBEAPVPNNKAYKIKKIDMTTYSKLTFTFIRAIIFHNVEK
PESITFGDGLCKEKQDLGIVFPYPLGGSDVDSDIPQVQEDALPAVSVLTG
FERLHLGKLLYLSQMSHVSAINTEYVRILPXDEDLFTKSSLRRLYLPAVSELYLY
TLVLSLQSLVHNAASIVEAIOQFPHDKKMKALVFSFKYPLATYGKMAQLQIHH
AAVEALIGSYLSIFEPAQOEKTKQDYDSMIDFASCEDSRDLRLKAWNLAQIHH
AQLSTNSIYVNVNVAOAPAPNSKVBPAVNVNSYLOHGLANLEETLFEDEGSAFG
APASIDGSEPTLOHLAVAAFPNSKVBPAVNVNSYLOHGLANLEETLFEDEGSAFG
NSPVCISLCSGCCPVCINTLTFYRLKDRPPVLYGSRBDPVYITITVFNELDLGNF
ASPRDKEDQNDQETPRYTYWLNQTLTELEAAGLVDSVADGAGGSGSNMLEKF
VRTSDIDSLIDATAAKAFINIMINNVNFKESIGSVHSDVIOYQNTYWAQAFCSLMLNL
YRSILITIDIALPISTVSESNPAGQYKRNEMLKHYQGLMTNFKSFIDKGVITG



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mpsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 23:27:13 1999; Maspar time 443.07 Seconds
972.040 Million cell updates/sec

Tabular output not generated.

Title: >US-09-240-410-1
Description: (1-2010) from US09240410.seq
Perfect Score: 2010
N.A. Sequence: 1 ATGACGGCTCTCCGCGCCG.....TGGTCACCTAGGGCTCCCG 2010
Comp: TACTGGGAGGAGGCGGGCC.....ACCGAGTATCCGCGAGGCG

Scoring table: TABLE default
Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 271905 seqs, 107135622 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

n-genesegs5
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40 41:part41 42:part42 43:part43
44:part44 45:part45 46:part46 47:part47 48:part48
49:part49 50:part50 51:part51 52:part52 53:part53
54:part54 55:part55 56:part56 57:part57 58:part58
59:part59 60:part60

Statistics: Mean 9.556; Variance 6.017; scale 1.588

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------|--------|----|--------|-----------------------|-----------|
| 1 | 47 | 2.3 | 190 | 32 | T76452 | Chymase antisense oli | 6.53e-12 |
| 2 | 44 | 2.2 | 178 | 32 | T76405 | Human endothelin-1 an | 3.05e-10 |
| 3 | 44 | 2.2 | 204 | 1 | N81164 | Base substituted E.co | 3.83e-09 |
| 4 | 42 | 2.1 | 91 | 9 | O51746 | Oligonucleotide probe | 1.08e-09 |
| 5 | 43 | 2.1 | 91 | 9 | O51746 | Oligonucleotide probe | 1.08e-09 |
| 6 | 43 | 2.1 | 204 | 1 | N81164 | Base substituted E.co | 1.08e-09 |
| 7 | 38 | 1.9 | 114 | 12 | O70467 | Generic DNA sequence | 5.46e-07 |
| 8 | 39 | 1.9 | 114 | 12 | O70469 | Generic DNA sequence | 5.46e-07 |
| 9 | 38 | 1.9 | 114 | 12 | O70465 | Generic DNA sequence | 5.46e-07 |

| | | | | | | | | |
|---|----|----|-----|------|----|--------|------------------------|----------|
| C | 10 | 38 | 1.9 | 114 | 12 | O70467 | Generic DNA sequence | 5.46e-07 |
| C | 11 | 38 | 1.9 | 144 | 32 | T76416 | Human endothelin rece | 5.46e-07 |
| C | 12 | 38 | 1.9 | 172 | 32 | T76363 | Human interleukin 8 a | 5.46e-07 |
| C | 13 | 36 | 1.8 | 89 | 32 | T76219 | Human IL5 antisense o | 1.85e-06 |
| C | 14 | 37 | 1.8 | 114 | 12 | O70469 | Generic DNA sequence | 1.85e-06 |
| C | 15 | 36 | 1.8 | 114 | 12 | O70468 | Generic DNA sequence | 6.18e-06 |
| C | 16 | 36 | 1.8 | 114 | 12 | O70472 | Generic DNA sequence | 6.18e-06 |
| C | 17 | 36 | 1.8 | 114 | 12 | O70470 | Generic DNA sequence | 6.18e-06 |
| C | 18 | 36 | 1.8 | 114 | 12 | O70465 | Generic DNA sequence | 6.18e-06 |
| C | 19 | 36 | 1.8 | 114 | 12 | O70471 | Generic DNA sequence | 6.18e-06 |
| C | 20 | 37 | 1.8 | 114 | 12 | O70466 | Generic DNA sequence | 1.85e-06 |
| C | 21 | 36 | 1.8 | 114 | 12 | O70468 | Generic DNA sequence | 6.18e-06 |
| C | 22 | 37 | 1.8 | 128 | 32 | T76233 | Human IL6 antisense o | 1.85e-06 |
| C | 23 | 37 | 1.8 | 168 | 32 | T76270 | Human MDNCF antisense | 6.71e-05 |
| C | 24 | 34 | 1.7 | 88 | 32 | T76110 | Human IL3 receptor an | 6.71e-05 |
| C | 25 | 35 | 1.7 | 114 | 12 | O70465 | Generic DNA sequence | 6.71e-05 |
| C | 26 | 34 | 1.7 | 114 | 12 | O70473 | Generic DNA sequence | 6.71e-05 |
| C | 27 | 34 | 1.7 | 264 | 32 | T76445 | Substance P receptor | 6.71e-05 |
| C | 28 | 32 | 1.6 | 39 | 7 | O51787 | Mixed oligonucleotide | 6.95e-04 |
| C | 29 | 33 | 1.6 | 88 | 32 | T76170 | Human IL3 receptor an | 6.95e-04 |
| C | 30 | 32 | 1.6 | 114 | 12 | O70471 | Generic DNA sequence | 6.95e-04 |
| C | 31 | 32 | 1.6 | 114 | 12 | O70470 | Generic DNA sequence | 6.95e-04 |
| C | 32 | 32 | 1.6 | 147 | 32 | T76234 | Human defensin 1 anti | 6.95e-04 |
| C | 33 | 33 | 1.6 | 160 | 47 | V48098 | Oligonucleotide Lp160 | 2.17e-04 |
| C | 34 | 33 | 1.6 | 162 | 32 | T76307 | Human RANTES antisense | 6.95e-04 |
| C | 35 | 32 | 1.6 | 201 | 2 | N70195 | Streptomyces protease | 6.95e-04 |
| C | 36 | 32 | 1.6 | 201 | 2 | N70194 | Signal portion of gen | 6.95e-04 |
| C | 37 | 32 | 1.6 | 250 | 32 | T76438 | Substance P antisense | 2.17e-04 |
| C | 38 | 33 | 1.6 | 317 | 32 | T76274 | Human neutrophil elas | 2.17e-04 |
| C | 39 | 31 | 1.5 | 100 | 32 | T76186 | Human IL4 receptor an | 2.19e-03 |
| C | 40 | 31 | 1.5 | 160 | 47 | V48098 | Oligonucleotide Lp160 | 2.19e-03 |
| C | 41 | 31 | 1.5 | 999 | 39 | T93598 | Elmeria tenella spore | 2.19e-03 |
| C | 42 | 31 | 1.5 | 1001 | 3 | N92578 | Sequence of the 1.7kb | 2.19e-03 |
| C | 43 | 31 | 1.5 | 1513 | 1 | O06160 | Chicken ovalbumin ups | 2.19e-03 |
| C | 44 | 31 | 1.5 | 3955 | 40 | V15181 | Human serrate 2 encod | 2.19e-03 |
| C | 45 | 31 | 1.5 | 4136 | 37 | T95368 | Human protein tyrosin | 2.19e-03 |

ALIGNMENTS

RESULT 1
ID T76452 standard; DNA: 190 BP.
AC T76452:
DE 16-SEP-1997 (first entry)
KW Chymase antisense oligonucleotide.
KW Asthma; airway epithelium; adenovirus free; cystic fibrosis;
KW Chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN MO9640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; 409306
PR 07-JUN-1995; US-474497
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger WJ, Nye JW;
DR WPI: 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenovirus-free antisense oligo:nucleotide to airway epithelium of
PT subject.
PS Example 5; Page 40; 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenovirus
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for chymase, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenovirus
CC from the antisense ON, its liberation upon adenovirus degradation is
CC prevented, thereby preventing adenovirus-induced bronchoconstriction
CC in patients with hyper-reactive airways.
SQ Sequence 190 BP; 1 A; 41 C; 58 G; 36 T;
Query Match 2.3%; Score 47; DB 3; Length 190;

```
Best Local Similarity 40.9%; Pred No.6-53e-12;
Matches 45; Conservative 36; Mismatches 28; Indels 1; Gaps 1

Db      24 bbggbbgbdbccctgttbtcbcbgccttcbg-.ctctggbbgcbdbbbgbbdbgbgcgcbg 82
       :|:::||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Oy      22 cctgccgcgccccacagcgaaccggcgcccgcgcncccttgccgccggcgctcttgaggcttt 81
                                     ||::|::|::|::|::|::|::|::|::|::|::|

Db      83 gggggygbbgbbbcbgcbgtcttcctccgbbgbygctgcttccttbgcbbbtgc 132
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy      82 ccgcttcgcgcttcgacctctgctctgctctgcttcggcgccggcgccgctccgc 131
                                     |::|::|::|::|::|::|::|::|::|::|::|

RESULT          2
ID             T76A05 standard; DNA: 178 BP.
AC             T76A05;
DC             15-SEP-1997 (first entry)
DE             Human endothelin-1 antisense oligonucleotide.
KW             Asthma; airway epithelium; adenosome free; cystic fibrosis;
               chronic obstructive pulmonary disease; bronchitis; ss.
OS             Synthetic.
PN             WO9640162-A1.
PD             19-DEC-1996.
PR             06-JUN-1996; U093306.
PF             07-JUN-1995; US-474497.
PA             (UYEC-) UNIV EAST CAROLINA.
PI             Metzger NJ, Nyece JW;
DR             WPL; 97-051871/05.
PT             Treatment of airway diseases such as asthma - by topically applying
PT             adenosome-free antisense oligo:nucleotide to airway epithelium of
PT             subject
PS             Claim 5: Page 38; 71pp; English.
CC             A method for treating airway disease in a subject has been produced,
CC             which involves the topical administration of an essentially adenome
CC             free antisense oligonucleotide (ON) to the airway epithelium of the
CC             subject. The present sequence is an antisense oligonucleotide specific
CC             for the human endothelin-1, targeted at the initiation codon. The
CC             method can be used to treat airway diseases such as cystic fibrosis,
CC             asthma, chronic obstructive pulmonary disease, bronchitis and other
CC             airway diseases characterised by an inflammatory response. By
CC             eliminating adenosome from the antisense ON, its liberation upon
CC             antisense degradation is prevented, thereby preventing adenosome-
CC             induced bronchoconstriction in patients with hyper-reactive airways.
SO             Sequence 178 BP; 0 A; 52 C; 46 G; 32 T;

Query Match           2.2% Score 44; DB 32; Length 178;
Best Local Similarity 33.1%; Pred. No. 3.05e-10;
Matches 39; Conservative 42; Mismatches 37; Indels 0; Gaps 0;

Db      53 bccbcbcgagctggtgcctbcgagcctbbgcbtgtctgtttctgbygactcttgcbgcbgc 112
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy      16 CCCGAGCTGCCGCCCGCCACGCGCAGCGCGCGCGCCGCCTCTGGCGCGCGCGCTTGATG 75
                                     ||::|::|::|::|::|::|::|::|::|::|::|

Db      113 bbhbcbgcbgbgbbdbbttcrtgbgcbdbbbttcbtctcgtbbabbhgaggtctbbbbb 170
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy      76 GGGCTTCCGCTGCGGCTGCGGCTCTGTCTGTCTGTGCGGCGCGCGCTTCGCCCC 133
                                     |::|::|::|::|::|::|::|::|::|::|::|

RESULT          3
ID             N81I64 standard; DNA: 204 BP.
AC             N81I64;
DC             08-NOV-1990 (first entry)
DE             Base substituted E.coli beta-galactosidase alpha-fragment.
KW             E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS             Escherichia coli.
FH             Key
FT             misc-feature
FT             Location/Qualifiers
FT             19..69
FT             /*tag= a
FT             /*function=multiple cloning site
FT             primer_bind
FT             187..204
FT             /*tag= b
FN             FP-285123-A.
FD             05-MAY-1988.
FE             30-MAR-1988. 105163.
```

[illegible]


```

Oy 23 GAGCGCCCGACGCGACCGGCGCGCGCCGCGCCTTCGAGCGCGGATTTGGAGGCTTC 82
Db 67 cgcgtgtgctcbbgbbgdtg-tgcttbccttcdbcbgagctcgbdbbctcbbgagct 125
Oy 83 GCGTCCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCGGCGCGCGCTCCGCCACGAGGCGACC 142
Db 126 gccbbgbbgagcgcgcgcgttgagcbb 155
Oy 143 TAAGAGCGGACCCCGCATTTCCGCGCT 172

RESULT 13
ID T76219 standard; DNA: 89 BP.
AC T76219.
DT 12-SEP-1997 (first entry)
DE Human IL5 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; interleukin; ss.
OS Synthetic.
PN W09640162.A1.
PD 19-DEC-1996.
PE 06-JUN-1996; U093306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger MJ, Nyce JW;
DR WPI: 97-051871/05.
PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS Claim 5: Page 30: 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human IL5, targeted at the initiation codon. The method can be
CC used to treat airway diseases such as cystic fibrosis, asthma, chronic
CC obstructive pulmonary disease, bronchitis and other airway diseases
CC characterised by an inflammatory response. By eliminating adenosine from
CC the antisense ON, its liberation upon antisense degradation is
CC prevented, thereby preventing adenosine-induced bronchoconstriction in
CC patients with hyper-reactive airways.
SQ Sequence 89 BP; 0 A; 20 C; 25 G; 18 T;

Query Match 1.8%; Score 36; DB 32; Length 89;
Best Local Similarity 46.4%; Pred. No. 6.18e-06;
Matches 26; Conservative 20; Mismatches 10; Indels 0; Gaps 0;

Db 21 ggcbbcbcbgcbgagcgtcbbgagctcbbgcbcbtcbtgcbbgagcctctcb 76
Oy 77 GCGTTCGCGTGCCTGCGGCTGCTGCTGCTGCTGCTGCGGCGCGCGCTCCGCC 132

RESULT 14
ID Q70469 standard; DNA: 114 BP.
AC Q70469.
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key
FT Location/Qualifiers
FT misc_feature 55..60
FT /tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN W09418318-A.
PD 18-AUG-1994.
PE 01-FEB-1994; U009977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.

```

| | |
|-----------------------|---|
| PA | (YUNC-) UNIV NORTH CAROLINA. |
| PI | Fowlkes DM, Kay BK; |
| PT | WPI: 94-279739/34. |
| PR | Identifying proteins or peptide(s) which bind a ligand - by |
| PS | screening a recombinant vector library expressing fusion proteins |
| PP | comprising a binding domain and an effector domain |
| PB | Disclosure; Page 35; 255pp; English. |
| CC | Q70465 is a generic DNA sequence used to generate random TSAR peptide |
| CC | This generic formula can be represented as follows: X(TGC)(NNB)10- |
| CC | (TGC)(NNB)6(NNB)2(TGC)(NNB)14(TGC)Y. x and y are flanking restriction |
| CC | sites (X is not the same as Y) that are not specified further. This |
| CC | sequence generates peptides that are cloverleaf in structure. Other |
| CC | generic sequences are shown in Q70465-68. Other specific peptides |
| CC | generated by these generic sequences are shown in R5150-34. TSARs are |
| CC | concatenated heterofunctional proteins or peptides, comprising at least |
| CC | two functional regions - a binding domain with affinity for a ligand and |
| CC | a second effector peptide portion that is chemically or biologically |
| CC | active.They may further comprise a linker peptide between the 2 domains. |
| CC | The oligonucleotides are also designed so that the expressed peptide |
| CC | contains 2 or 4 cysteine residues positioned in, or flanking, the |
| CC | unpredicted or variant residues. These residues confer some degree of |
| CC | conformational rigidity to the peptides. The TSARs or compsns. comprising |
| CC | a TSAR binding domain can be used in vivo to deliver a chemically or |
| CC | biologically active moiety, eg. metal ion, radioisotope, peptide, toxin |
| CC | or enzyme, to the specific target or on the cell. They can also replace |
| CC | the function of macromolecules, eg. monoclonal or polyclonal antibodies |
| CC | and therefore circumvent the need for complex methods of hybridoma |
| CC | formation or in vivo antibody production. The TSARs are easily |
| CC | characterised and have designed activity allowing direct and rapid |
| CC | detection in a screening process. |
| SQ | Sequence 114 BP; 0 A; 4 C; 4 G; 4 T; |
|
 | |
| Query Match | 1.8%; Score 37; DB 12; Length 114; |
| Best Local Similarity | 8.3%; Pred. No. 1.85e-06; |
| Matches | 9; Conservative 30; Mismatches 70; Indels 0; Gaps 0; |
|
 | |
| Dc | 6 bmbnbmbnbnbnbnbnbnbnbnbtgcmbnbnbnbnbnbnbnbnbnbnbnbn 65 |
| Qy | 1858 TACTTCGCGAGGCTCAGACGTGGCAGCTGCGCCCGAGGAGCATCATGCCGAGCAC 1917 |
| Dc | 66 btgcnbtgc 114 |
| Qy | 1918 CTGTGGTGATCATGCTGTGCCCTGGCGCCGCTCTCTGTGGGGGTGC 1966 |
|
 | |
| RESULT | 15 |
| ID | Q70468 standard; DNA; 114 BP. |
| AC | Q70468; |
| DT | 05-APR-1995 (first entry) |
| DE | Generic DNA sequence to generate a random TSAR peptide library. |
| KM | TSAR: totally synthetic affinity reagent; synthetic; binding domain; |
| KW | effector domain; concatenated heterofunctional protein; linker; |
| KW | direct; rapid; detection; screening; treatment; genetic; ss. |
| OS | Synthetic. |
| FH | Key |
| FT | misc_feature |
| FT | Location/Qualifiers |
| FT | 55..60 |
| FT | /tag= a |
| FT | /note= "this sequence represents 'Z'; Z can be a |
| FT | sequence of 6, 9 or 12 nucleotides (see |
| FT | comments)" |
|
 | |
| PN | WC09418318-A. |
| PD | 18-AUG-1994. |
| PE | 01-FEB-1994; U00977. |
| PR | 01-FEB-1993; US-013416. |
| PR | 30-DEC-1993; US-176500. |
| PR | 31-JAN-1994; US-189331. |
| PA | (YUNC-) UNIV NORTH CAROLINA. |
| PI | Fowlkes DM, Kay BK; |
| PT | WPI: 94-279739/34. |
| PR | Identifying proteins or peptide(s) which bind a ligand - by |
| PT | screening a recombinant vector library expressing fusion proteins |
| PT | comprising a binding domain and an effector domain |

PS Disclosure: Page 35; 255bp; English.

CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
 CC Synthetic Affinity Reagents) peptides. This generic formula can also be
 CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X
 CC and Y are flanking restriction sites (X is not the same as Y) that are
 CC not specified further. Other generic sequences are shown in Q70466-68.
 CC Other specific peptides generated by these generic sequences are shown in
 CC R5151-54. TSARs are concatenated heterofunctional proteins or peptides,
 CC comprising at least two functional regions - a binding domain with
 CC affinity for a ligand and a second effector peptide portion that is
 CC chemically or biologically active. They may further comprise a linker
 CC peptide between the 2 domains. The oligonucleotides are also designed so
 CC that the expressed peptide contains 2 or 4 cysteine residues positioned
 CC in, or flanking, the unpredicted or variant residues. These residues
 CC confer some degree of conformational rigidity to the peptides. The TSARs
 CC or compns. comprising a TSAR binding domain can be used in vivo to
 CC deliver a chemically or biologically active moiety, eg. metal ion,
 CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
 CC cell. They can also replace the function of macromolecules, eg.
 CC monoclonal or polyclonal antibodies and therefore circumvent the need
 CC for complex methods of hybridoma formation or in vivo antibody
 CC production. The TSARs are easily characterized and have designed activity
 CC allowing direct and rapid detection in a screening process.

SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.8%; Score 36; DB 12; Length 114;
 Best Local Similarity 4.5%; Pred. No. 6.18e-06;
 Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnnnnnnnn 62
 OY 22 CGTCCGCCGCCAGCGACCGCGCGCCGCTCCCTGGCGCCGCGCTGGGGCCTT 81

Db 63 bnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnbdbnb 114
 OY 82 CCGCTGGCGCTGCGCTGCTGCTGCTGCGCGCGCGCGCTCCGCC 133

Search completed: Tue Aug 3 23:35:37 1999
 Job time : 504 secs.

This Page Blank (uspto)

| Query Match | 19.2% | Score 386 | DB 34 | Length 424 |
|-----------------------|----------------|---|----------|------------|
| Best Local Similarity | 97.6% | Pred. No. 0.00e+00 | | |
| Matches 415 | Conservative 0 | Mismatches 6 | Indels 4 | Gaps 4 |
| Db | 2 | TGAACTTCTCCCAACCTGTGCTCCCTCCCTGCACAACTGGGCGCACAGGACATGAGAGG | 61 | |
| Cp | 825 | TGAACCTTCCCAACCTGTGCTCCCTCCCTGCACAACTGGGCGCACAGGACATGAGAGG | 766 | |
| Db | 62 | AGCCTCAGAGATCTTGTACAGATTGTCTCTTGGGAAGTAGTAGATCTTGTACGTGA | 121 | |
| Cp | 765 | AGCCTCAGAGATCTTGTACAGATTGTCTCTTGGGAAGTAGTAGATCTTGTACGTGA | 706 | |
| Db | 122 | AGCCTGCTTGTGTGTCACAGATGTGTGCTTGTATGAACTGTGTGTTCATGACATATC | 161 | |
| Cp | 705 | AGCCTGCTTGTGTGTCACAGATGTGTGCTTGTATGAACTGTGTGTTCATGACATATC | 646 | |
| Db | 182 | ACTGCTGTACAGACTACTCTGCCCCGGATGCGGCGGAACGAGAGATCTTCCATGTA | 241 | |
| Cp | 645 | ACTGCTGTACAGACTACTCTGCCCCGGATGCGGCGGAACGAGAGATCTTCCATGTA | 586 | |
| Db | 242 | TTTCCTGCTTCCGGATGTGTGAATACACTGTCTCCCTTCAAAACGAACAGGGACGTTT | 301 | |
| Cp | 585 | TTTCCTGCTTCCGGATGTGTGAATACACTGTCTCCCTTCAAAACGAACAGGGACGTTT | 528 | |
| Db | 302 | TGCTTCGGGCTGAAGGGGGCGTCTCTAC-TCTGCGAAGTGGCCACACAGTGCATNTC | 360 | |
| Cp | 527 | TGCTTCGGGCTGAAGGGGGCGTCTCTAC-TCTGCGAAGTGGCCACACAGTGCATNTC | 468 | |
| Db | 361 | ACCAGGTTCCACAGACTGGGGTGCCTGGTGTGTCACAGGC-ACGACCCCTCCTACTC | 419 | |
| Cp | 467 | ACCAGGTTCCACAGACTGGGGTGCCTGGTGTGTCACAGGC-ACGACCCCTCCTACTC | 408 | |
| Db | 420 | CGCCT 424 | | |
| Cp | 407 | CGCCT 403 | | |

| RESULT | 5 | 446 bp | mRNA | EST | 20-JUN-1995 |
|------------|--|-----------|------|-----|-------------|
| LOCUS | H02902 | | | | |
| DEFINITION | y139f01.s1 Soares placenta ND2HP Homo sapiens cDNA clone | | | | |
| ACCESSION | H02902 | | | | |
| KEYWORDS | g865835 | | | | |
| VERSION | H02902.1 | GI:865835 | | | |
| KEYWORDS | EST | | | | |
| SOURCE | human | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 446)
Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maria, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R. | | | | |
| TITLE | The Washu-Merck EST-Project | | | | |
| JOURNAL | Unpublished (1995) | | | | |
| COMMENT | Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1575
High quality sequence stops: 196
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1575 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 196. | | | | |

| FEATURES | Location/Qualifiers |
|-------------|--|
| SOURCE | 1..446
/organism="Homo sapiens"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGCAAGATTTCGGCGCCGACAGATTTTCTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo." |
| BASE COUNT | 93 a 115 c 136 g 93 t 9 others |
| ORIGIN | |
| Query Match | 18.2% Score 365; DB 30; Length 446;
Best Local Similarity 95.4%; Pred. NO. 0.00e+00; |
| Matches | 396; Conservative 0; Mismatch 15; Indels 4; Gaps 4; |
| Db | 1 GCCGCCTGAAGAGCCATGTAGGAGCAGACC GGGTGGACTTTGGCCACGATGACGCCGAC 60 |
| OY | 166 GCCCTCTGAAAGCCCATGTAGGCGACGACCGGTGCACTTTGGCCACGATGACGCCGAC 225 |
| Db | 61 ACGGTCTTTTCCACGACCCAGGACGCTCTCTGTGTGGTGAGGAGCAGTGGCAAGTTC 120 |
| OY | 226 ACGGTCTTTTCCACGACCCAGGACGCTCTCTGTGTGGTGAGGAGCAGTGGCAAGTTC 285 |
| Db | 121 TACCTCTTTGACTTCCCCCGAAGGSCAAGAAGCATCTGTGGCACAGGTGAATATCGCTCC 180 |
| OY | 286 TACCTCTTTGACTTCCCCCGAAGGSCAAGAAGCATCTGTGGCACAGGTGAATATCGCTCC 345 |
| Db | 181 ACAAGAGGCTCTCTGTGATACGGGAGCTGCGACAAGTAACATCATCTCCCTGGAGAGG 240 |
| OY | 346 ACAAGAGGCTCTCTGTGATACGGGAGCTGCGACAAGTAACATCATCTCTCTGGAGAGG 405 |
| Db | 241 CGGANTGAGGGGCTGCTGCGCTGTG- ACCAACGCCCGGACCCGACCTGCTGGAACCTG 299 |
| OY | 406 CGGAGTGAAGGGCTGCTGCGCTGTGCGACCAACGCCCGGACCCGACCTGCTGGAACCTG 465 |
| Db | 300 GGTGAATGACCTNTGGTCAC- CTTGGCGAGATAGAGGCTTAGCCCCCTTCAGCCNCA 358 |
| OY | 466 G-TAAATGGCACTGTGGTCCACTTGGCGAGATGAGAGGCTTAGCCCCCTTCAGCCGGA 524 |
| Db | 359 CGAGACTTCTCTGTTCTNTTTTMAAGGGAGCAGATTTATTCANCATCCNCAA 413 |
| OY | 525 CGAGACTTCTCTGTTCTNTTTTMAAGGGAGCAGATTTATTCANCATCCGGA 578 |
| RESULT | 6 |
| LOCUS | H03806 434 bp mRNA EST 20-JUN-1995 |
| DEFINITION | YJ39F01.r1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:151129 5' similar to SP:A49069 A49069 COLLAPSPIN -, mRNA sequence. |
| ACCESSION | H03806 |
| NID | 9666739 |
| VERSION | H03806.1 GI:866739 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 434)
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Morris,M., Parsons,J., Rifkin,I., Rohlfing,T., Soares,M., Tan,F. |


```

Oy 247 GGCAGCTCTCTGTGTGGTGGAGAGACGTGCGAAGTCTACCTCTTGTGACTTCCCGAG 306
      /clone.lib="NCI_CGAP_GCH1"
      /tissue-type="germinal center B cell"
      /lab-host="DH10B"
Db 121 GGCAGAGCATCTGTGGCGACGAGTGAATATGGCTCCCAAGAGGGTCTCTTGAT 180
      |||||||
Oy 307 GGCAGAGCATCTGTGGCGACGAGTGAATATGGCTCCCAAGAGGGTCTCTTGAT 366
      |||||||
Db 181 AACGGGACTGCGAGACATACATCTCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 240
      |||||||
Oy 367 AACGGGACTGCGAGACATACATCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 426
      |||||||
Db 241 TGTGGACCAACAGCCCGGACCCAGCTGTGTGAACCTGTGTATGTGCTGTGGCA 300
      |||||||
Oy 427 TGTGGACCAACAGCCCGGACCCAGCTGTGTGAACCTGTGTATGTGCTGTGGCA 486
      |||||||
Db 301 CTGGCGAGATGAGAGCTACGCCCCCTTCAAGCCG 336
      |||||||
Oy 487 CTGGCGAGATGAGAGCTACGCCCCCTTCAAGCCG 522
      |||||||

RESULT 8
LOCUS AA856724 332 bp mRNA EST 09-MAR-1998
DEFINITION od54c01.s1 NCI_CGAP_GCH1 Homo sapiens cDNA clone IMAGE:1371744 3'
      similar to TR:064906 Q84906 SIMILAR TO GENBANK ACCESSION NUMBER
      L26081.1, mRNA sequence.
ACCESSION AA856724
NID 92945026
VERSION AA856724.1 GI:2945026
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
      Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 332)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
      Tumor Gene Index
      Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2151919.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
      Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
      Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
      clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 323.
Location/Qualifiers
1. 332
/organism="Homo sapiens"
/Note="Vector: pT733-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I-Oligo(dT) primer.
[5'-TGTACCAATCTGAGTGGAGCGCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT733 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/map="5: 5q23.3-5q31.1"
/clone="IMAGE:1371744"

Oy 247 GGCAGCTCTCTGTGTGGTGGAGAGACGTGCGAAGTCTACCTCTTGTGACTTCCCGAG 306
      /clone.lib="NCI_CGAP_GCH1"
      /tissue-type="germinal center B cell"
      /lab-host="DH10B"
Db 121 GGCAGAGCATCTGTGGCGACGAGTGAATATGGCTCCCAAGAGGGTCTCTTGAT 180
      |||||||
Oy 307 GGCAGAGCATCTGTGGCGACGAGTGAATATGGCTCCCAAGAGGGTCTCTTGAT 366
      |||||||
Db 181 AACGGGACTGCGAGACATACATCTCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 240
      |||||||
Oy 367 AACGGGACTGCGAGACATACATCTCTCTGAGAGAGCGGAGTGGGCTGTGGCC 426
      |||||||
Db 241 TGTGGACCAACAGCCCGGACCCAGCTGTGTGAACCTGTGTATGTGCTGTGGCA 300
      |||||||
Oy 427 TGTGGACCAACAGCCCGGACCCAGCTGTGTGAACCTGTGTATGTGCTGTGGCA 486
      |||||||
Db 301 CTGGCGAGATGAGAGCTACGCCCCCTTCAAGCCG 336
      |||||||
Oy 487 CTGGCGAGATGAGAGCTACGCCCCCTTCAAGCCG 522
      |||||||

Query Match 15.8% Score 318; DB 20; Length 332;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 331; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
ORIGIN
BASE COUNT 66 a 95 c 109 g 62 t

Db 1 GGACTTTGGCCAGACTGAGACCGCAGACAGGTGCTTTTCCAGAGCAGGACGCTCTGT 60
      |||||||
Oy 201 GGACTTTGGCCAGACTGAGACCGCAGACAGGTGCTTTTCCAGAGCAGGACGCTCTGT 260
      |||||||
Db 61 GTGGGTGGAGAGACCTGCGAGAGTCTACTCTTTACTTCCCGAGGCGCAGAGCATC 120
      |||||||
Oy 261 GTGGGTGGAGAGACCTGCGAGAGTCTACTCTTTACTTCCCGAGGCGCAGAGCATC 320
      |||||||
Db 121 TGTGGCAGCGTGAATATCGGCTCCACAAGAGGCTGCTGTGATAGCGGAGTGGCA 180
      |||||||
Oy 321 TGTGGCAGCGTGAATATCGGCTCCACAAGAGGCTGCTGTGATAGCGGAGTGGCA 380
      |||||||
Db 181 GAATACATCACTCTCTCTGAGAGAGCGAGTGAAGGCTGTGGCTGTGGCACCACGC 240
      |||||||
Oy 381 GAATACATCACTCTCTCTGAGAGAGCGAGTGAAGGCTGTGGCTGTGGCACCACGC 440
      |||||||
Db 241 ACGGCACCCAGCTCTCTGAGAGTGAATGACACTGTGTGC-ACITGGCGAGATGAG 299
      |||||||
Oy 441 CCGGCACCCAGCTCTCTGAGAGTGAATGACACTGTGTGC-ACITGGCGAGATGAG 500
      |||||||
Db 300 AGCTACGCCCC-TCAGCCCGGAGAGAGACTCC 332
      |||||||
Oy 501 AGCTACGCCCCCTTCAAGCCGAGAGACTCC 534
      |||||||

RESULT 9
LOCUS AA968218 422 bp mRNA EST 19-MAY-1998
DEFINITION uh14f09.r1 Soares mouse hypothalamus NMH9 Mus musculus cDNA clone
      IMAGE:1617929 5' similar to TR:064906 Q84906 SIMILAR TO GENBANK
      ACCESSION NUMBER L26081.1, mRNA sequence.
ACCESSION AA968218
NID 93142111
VERSION AA968218.1 GI:3142111
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 422)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
      Getzel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
      Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
      Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
      Waterston,R.
      The Mashu-HMI Mouse EST Project
      Unpublished (1996)
JOURNAL On Jan 17, 1998 this sequence version replaced gi:2043681.
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:957229
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 405.
Location/Qualifiers
1. 422
/organism="Mus musculus"

```

```

/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTATCCATCTGAGTGTGGAGCGCCGCAAGGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Wolfgang Liedtke. Library went through
two rounds of normalization, and was constructed by Bento
Soares and M.Fatima Ronaldo."
/db_xref="taxon:10090"
/clone="IMAGE:1617923"
/clone_1ib-"Soares mouse hypothalamus NMHy"
/tissue_type="hypothalamus"
/lab_host="DH10B"

```

| Query Match | Similarity | Score | DB | Length |
|-------------|------------|--------------------|----|---------------------------------|
| Best Local | 87.4% | Pred. No. 0.00e+00 | | |
| Matches | 365 | Conservative | 0 | Mismatches 53; Indels 0; Gaps 0 |

| Accession | LOCUS | EST | CDNA | Protein |
|-----------|-------|---|------|---------|
| Db | 1 | TGAGAGATATACAGGGCTATAGCGCTTTCTCCAAACCCCTGGAAGTACTAGCTGTCTGC | 60 | |
| Qy | 946 | TGAGAGGAGACACAGGGCTATAGTGTCTTCTCCAAACCCCTGGAAGTACTAGCGGTCTGT | 1005 | |
| Db | 61 | GTTGATTCGCTTGGTGACATTGACAGAGCTTCCGTACCTATCGCTCAAGGCTACCAAC | 120 | |
| Qy | 1006 | GTTGATTCCTTCGGTGACATTGACAGAGCTTCCGTACCTCTCACTCAAGGGCTACCAAC | 1065 | |
| Db | 121 | ATGGGCGCTTCCCAACCCCTCGAGCTGGCATGTGCGCTCCCAAAAAAGAGCAGCCATACCA | 180 | |
| Qy | 1066 | TCAAGCCTTCCCAACCCCGGGCTGGCAAGTGCCTCCAGACACAGCCGATACCCACA | 1125 | |
| Db | 181 | GAACCTTCCAGAGTACTGATAGTCACCCAGAGGTGGCTCAGAGGGTGAACCTATGGGG | 240 | |
| Qy | 1126 | GAGACTTCCAGAGTGGCTGACCGCTACCCAGAGGTGGCGCAGAGGGTGAAGCCATGGGG | 1185 | |
| Db | 241 | CCACTGAAGACACCATTTGTTCCATTTCTAAGTACCATTTACCGAAGTGTGCTTACCCGC | 300 | |
| Qy | 1186 | CCTTGAAGAGCGCATTTTCTCCACTTAATAACCACTACCGAAGTGGCGCTCACCGC | 1245 | |
| Db | 301 | ATGCAAGCCAGCAATGAGAGAGACCTTCATGCTTTATCAACACAGACAGGAGGCACC | 360 | |
| Qy | 1246 | ATGCAAGCCAGCCAGCGGGAGACACTTTATGTGCTTTACCTAATACAGACAGGAGCACT | 1305 | |
| Db | 361 | ATTCAACAAGTGTTGGTAATCAAGGGGACGAGCACTATGCTTTCATATATATATGAG | 420 | |
| Qy | 1306 | ATCCACAAGTGTTGGTAATCAAGGGGAGCAGAGCAGACACTTGCCTTCAACATCATGGAG | 1365 | |
| Db | 421 | AT 422 | | |
| Qy | 1366 | AT 1367 | | |

| Result | 10 | LOCUS | 344 bp | mRNA | EST | 28-Apr-1995 |
|------------|----|--|--------|------|-----|-------------|
| DEFINITION | | YH9803.f1 Soares Placenta ND2HP Homo sapiens cDNA clone | | | | |
| | | IMAGE:135941.5' similar to SP:VA59_VACC P21062 PROTEIN ; mRNA | | | | |
| | | sequence. | | | | |
| ACCESSION | | R33537 | | | | |
| NID | | 9789395 | | | | |
| VERSION | | R33537.1 | | | | |
| KEYWORDS | | EST. | | | | |
| SOURCE | | human. | | | | |
| ORGANISM | | Homo sapiens | | | | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | | | |
| | | Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| REFERENCE | | 1 (bases 1 to 344) | | | | |
| AUTHORS | | Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., | | | | |
| | | Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maitra, M., | | | | |
| | | Patson, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., | | | | |
| | | Tsvetakis, E., Waterston, R., Williamson, A., Wohlmann, P. and | | | | |

| FEATURES | TITLE |
|----------|--|
| Source | Wilson, R.
The Washu-Merck EST Project
(Unpublished (1995)) |
| | Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1078 |
| | High quality sequence stops: 275
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarly on wrong strand
Insert Length: 1078 Std Error: 0.00
Seq primer: M13Rpl |
| | High quality sequence stop: 275.
Location/Qualifiers
1..344 |

| BASE COUNT | 74 a | 87 c | 106 g | 75 t | 2 others |
|---------------------------|--|------|-------|------|----------|
| ORIGIN | | | | | |
| Query Match | 15.5%; Score 312; DB 30; Length 344; | | | | |
| Best Local Similarity | 98.0%; Pred. No. 0.00e+00; | | | | |
| Matches 337; Conservative | 0; Mismatches 3; Indels 4; Gaps 4; | | | | |
| Db | 1 GCTGCTGCTCTTGGGAGGACACTTGCACGCGCCCGGGGTTGGGAAGGCTTGATGCTGATGCCCT 60 | | | | |
| Cp | 1114 GCTGCTGCTCTTGGGAGGACACTTGCACGCGCGGGGTTGGGAAGGCTTGATGCTGATGCCCT 1055 | | | | |
| Db | 61 TGAGTGAGGAGGTACGGAAGACCTTGTAATGTACACCGAGGAATACACAGAGCGGCTG 120 | | | | |
| Cp | 1054 TGAGTGAGGAGGTACGGAAGACCTTGTCATGTATCACCAGAGGAATACACAGAGCGGCTG 995 | | | | |
| Db | 121 AGTACTTCCAGGGGTTGGAGAAACACCATAGACCCCTGGTGTCCCTCCACTGCGCGCTGG 180 | | | | |
| Cp | 994 AGTACTTCCAGGGGTTGGAGAAACACCATAGACCCCTGGTGTCCCTCCACTGCGCGCTGG 935 | | | | |
| Db | 181 GGTCAAGGAGACGAGGAAGACGCTTTCACGACCTTTGAAGTCTTGTGTTGGGAGACATCAC 240 | | | | |
| Cp | 934 GGTCAAGGAGACGAGGAAGACGCTTTCACGACCTTTGAAGTCTTGTGTTGGGAGACATCAC 875 | | | | |
| Db | 241 TGCATTACAGCATGGCTTTTCAGAAAAGTGTTCACCTTGGAGAGTACAGATGAACCTTT 300 | | | | |
| Cp | 874 TGCAT-ACCAGCATGGCTTT-CAGAAAAGTGTTCACCTT-GGAGACTGCAGATGAACCTTT 818 | | | | |
| Db | 301 CCCACCCCTGGTCCNCCTTGCACAAATTNGGCCACACGGGACAC 344 | | | | |
| Cp | 817 CCCACCCCTGGTCCNCCTT-GCACAATAGTGGCCACACGGGACAC 775 | | | | |
| RESULT | 11 | | | | |
| LOCUS | R33439 309 bp mRNA EST 28-APR-1995 | | | | |

| | | | | | | | | |
|-----------------------|---|---------------|-----------|---------|----------|--|--|--|
| DEFINITION | U79603.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
IMAGE:159941_3 , mRNA sequence. | | | | | | | |
| ACCESSION | R33439 | | | | | | | |
| NID | 9789297 | | | | | | | |
| VERSION | R33439.1 | GI:789297 | | | | | | |
| KEYWORDS | EST. | | | | | | | |
| SOURCE | human. | | | | | | | |
| ORGANISM | Homo sapiens | | | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 309)
Hiller,L., Clark,N., Dubaque,T., Elliston,K., Hawkins,M.,
Holman,J., Hultman,M., Kucaba,T., Le'M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaslis,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R. | | | | | | | |
| TITLE | The WashU-Merck EST Project | | | | | | | |
| JOURNAL | Unpublished (1995) | | | | | | | |
| COMMENT | <p>Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1078</p> <p>High quality sequence stops: 273
Source: IMAGE consortium, LLNL</p> <p>This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: poly not found
Insert Length: 1078 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 273.</p> <p>Location/Qualifiers
1 . 309
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pT73D (Pharmacia) with
modified polylinker; Site1: Not I; Site2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
AACGGAAGAATTTGCCGGCCCGCAGAAATTCTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Benito Soares and M.Fatima Bonaldo.
/db_xref="GeneBank:U79603"
/map="10 q11.2-q22.1"
/clone_id="IMAGE:135941"
/clone_lib="Soares Placenta Nb2HP"
/sex="Female"
/dev_stage="Placenta obtained at birth (full term)"
/_lab_host="DH10B (ampicillin resistant)"</p> | | | | | | | |
| BASE COUNT | 58 a | 86 c | 104 g | 58 t | 3 others | | | |
| ORIGIN | | | | | | | | |
| Query Match | 14.2%; Score 286; DB 30; Length 309; | | | | | | | |
| Best Local Similarity | 97.7%; Pred. No. 0.00e+00; | | | | | | | |
| Matches | 295; Conservative | 0; Mismatches | 6; Indels | 1; Gaps | 1; | | | |
| Db | 8 CCGCCAGAGGCACCCTAAGAGCGGATTCANNNTCTTGCGCGTGTGMAAGGCCATGTAG | 67 | | | | | | |
| QY | 128 CCGGCCAGGCGCACCACTTAAGAGCGGACCCGACCATCTTCGCGTGTGMAAGGCCATGTAG | 187 | | | | | | |
| Db | 68 GGCGAGACCGGGGTGACTTTGGCCAGACTAGACCGGACACGGGTCTTTCCAGAGGCCAG | 127 | | | | | | |
| OY | 188 GGCGAGACCGGGGTGACTTTGGCCAGACTAGACCGGACACGGGTCTTTCCAGAGGCCAG | 247 | | | | | | |
| Db | 128 GCACCTCCTCTGTGTGGGTGGGAGAGAGTAGTGACAAGTCACTTGTGATTCGCCGAG | 187 | | | | | | |
| OY | 248 GCACCTCCTCTGTGTGGGTGGGAGAGAGTAGTGACAAGTCACTTGTGATTCGCCGAG | 307 | | | | | | |

| Db | 188 | GCAGAACGCATCTGTGCGCACGGTGTGAATATCGGCTCCACAAAGGGGCTGTCTGGATA | 247 |
|------------|---|--|-------------|
| Qy | 308 | GCAGAACGCATCTGTGCGCACGGTGTGAATATCGGCTCCACAAAGGGGCTGTCTGGATA <td>367</td> | 367 |
| Db | 248 | ACGGGAGCTGCGAGACTACATCACTCTCTGGAGAGGCGGAGTGAAGGGCTGCTGGCC | 307 |
| Qy | 368 | AGCGGAGCTGCGAGACTACATCACTCTCTGG-AGAGGCGGAGTGAAGGGCTGCTGGCC | 426 |
| Db | 308 | TG 309 | |
| Qy | 427 | TG 428 | |
| RESULT | 12 | | |
| LOCUS | AA260340 | 347 bp | MRNA |
| DEFINITION | va91b02.r1 Soares mouse 3MNE12.5 Mus musculus cDNA clone | | 18-MAR-1997 |
| | IMAGE:746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK | | |
| | ACCESSION NUMBER L26081. ; mRNA sequence. | | |
| ACCESSION | AA260340 | | |
| NID | g1896824 | | |
| VERSION | AA260340.1 | GI:1896824 | |
| KEYWORDS | EST. | | |
| SOURCE | house mouse. | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| AUTHORS | 1 (bases 1 to 347)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheinberg,R., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R. | | |
| TITLE | The WashU-HMI Mouse EST Project | | |
| JOURNAL | Unpublished (1996) | | |
| COMMENT | On Sep 12, 1996 this sequence version replaced gi:1404539. | | |
| | Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:45747
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 296.
Location/Qualifiers
1. 347
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: whole fetus; Vector: p773p-Pac (pharmacia)
with a modified polylinker; Site.1: Not I; Site.2: Eco RI
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' TGTTACCAATCTAGTCGAGCGGCGCCCTATTTTCTTTTCTTTT
3'] on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
Library went through one round of normalized p773 vector.
constructed by Bento Soares and M. Fatima Bonaldo. " | | |
| | /db_xref="taxon:10090"
/clone_id="Soares mouse 3MNE12.5"
/seq="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B" | | |
| BASE COUNT | 75 a | 102 c | 90 g 80 t |
| ORIGIN | | | |

Query Match 13.3%; Score 267; DB 10; Length 347;
 Best Local Similarity 88.7%; Pred. No. 0.00e+00;
 Matches 306; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Db 1 GAGAGACACACCTGACACAGACCCGAGGCTCTCTCAATGATCCCGAGTACCCAGT 60
 |||||
 QY 734 GAGAGACACACCTGACACAGATCTCTCTCAATGATCCCGAGTACCCAGT 793
 |||||
 Db 61 TGTCACAGGGGACACAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
 |||||
 QY 794 TGTCACAGGGGACACAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 853
 |||||
 Db 121 TGAAGGACATGT 180
 |||||
 QY 854 TGAAGGACATGT 913
 |||||
 Db 181 ATGCTTCTGT 240
 |||||
 QY 914 ACCTTCTGT 973
 |||||
 Db 241 TCTCCACCCCTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 |||||
 QY 974 TCTCCACCCCTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033
 |||||
 Db 301 TCTCCACCCCTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
 |||||
 QY 1034 TCTCCACCCCTGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
 |||||

RESULT 13
 LOCUS AA747293 157 bp mRNA EST 17-MAR-1999
 DEFINITION nx88a10.s1 NCI-CGAP-GCB1 Homo sapiens cDNA clone IMAGE:1269306 3'
 similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
 L26081.1, mRNA sequence.
 ACCESSION AA747293
 NID 92787251
 VERSION AA747293.1 GI:2787251
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 157)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:693526.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bdrp/image/image.html

Insert Length: 825 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

1..157
 /organism="Homo sapiens"
 /note="Vector: pT7TD-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, 19D-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAGTGGAGCGCGCGCTCATTTTCTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /map="1p: 5"
 /clone="IMAGE:1269306"
 /clone_lib="NCI-CGAP-GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"

BASE COUNT 33 a 43 c 54 g 26 t 1 others
 ORIGIN

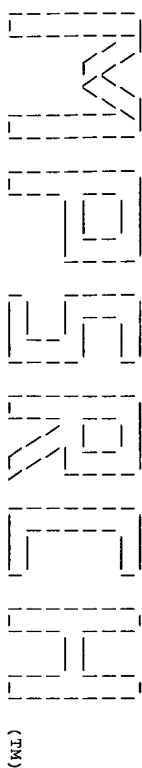
Query Match 6.4%; Score 129; DB 17; Length 157;
 Best Local Similarity 94.6%; Pred. No. 3.51e-181;
 Matches 141; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 Db 5 GGCAGAGACGATCTGTGCGCAGGATGATGCGCTCCACAAAGGGCTCTGTGTGAT 64
 |||||
 QY 307 GGCAGAGACGATCTGTGCGCAGGATGATGCGCTCCACAAAGGGCTCTGTGTGAT 366
 |||||
 Db 65 AAGCGGAGCTGCGAGACTACATCATCTCTCTGAGAGGGGAGTGGGCTGTGGC 124
 |||||
 QY 367 AAGCGGAGCTGCGAGACTACATCATCTCTCTGAGAGGGGAGTGGGCTGTGGC 426
 |||||
 Db 125 TGTGG-ACCAAGCGCCGCGACCCAGCTG 152
 |||||
 QY 427 TGTGGACCAAGCGCCGCGACCCAGCTG 455
 |||||

RESULT 14
 LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
 DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPLI cDNA library Oryza sativa
 cDNA clone 97SN1787, mRNA sequence.
 ACCESSION AA754459
 NID 92801165
 VERSION AA754459.1 GI:2801165
 KEYWORDS EST.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
 Poaceae; Oryza.
 REFERENCE 1 (bases 1 to 252)
 Naim, B.H., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,
 Kim, W.T., Kim, W.Y., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,
 Lee, M.C., and Eun, M.Y.
 Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
 Unpublished (1998)
 JOURNAL On Jan 14, 1998 this sequence version replaced gi:1797457.
 COMMENT

Contact: Eun M. Y.
 Department of Cytogenetics
 National Inst. of Agril. Sci. and Tech, RDA
 Suwon, Kyunggi-do, Korea
 Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeun@suno.asti.re.kr
 Submitted by Beek Hie Naim, Dept of Biological Science, Myongji
 University, Yongin, Korea. 449-728 bnamh@bserver.myongji.ac.kr
 Seq primer: M13 Reverse Primer.
 Location/Qualifiers

FEATURES

1..252
 /organism="Oryza sativa"
 /cultivar="Milyang23"
 /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
 XhoI; Directional cDNA library inserted into lambda ZAPRI
 vector at 5' end with EcoRI and 3' end with Xho I site."
 /db_xref="taxon:4530"



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 4 01:06:00 1999; Maspar time 1236.57 Seconds
1349.161 Million cell updates/sec

Tabular output not generated.

Title: >US-09-240-410-3
Description: (1-712) from US09240410.seq
Perfect Score: 708
N.A. Sequence: 1 CCGCGAGGCGGCGGCGGCA.....CCAGGGTGGGGAAGTTCAN 712
Comp: GCGCGAGGCGGCGGCGGCGG.....GGTCCACCCCTTCAAGTN

Scoring table: TABLE default

Gap 6

Mmatch SMD : Dbase 0: Query 0

Searched: 2883791 segs, 1171580779 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est58
1:em_est10 2:em_est11 3:em_est17 4:em_est18 5:em_est2

Database: 6:em_est17 7:em_gss1
genbank-est111

8:gb_est1 9:gb_est10 10:gb_est11 11:gb_est12 12:gb_est13
13:gb_est14 14:gb_est15 15:gb_est16 16:gb_est17
17:gb_est18 18:gb_est19 19:gb_est20 20:gb_est21
21:gb_est22 22:gb_est23 23:gb_est24 24:gb_est25
25:gb_est26 26:gb_est27 27:gb_est28 28:gb_est29
29:gb_est30 30:gb_est31 31:gb_est32 32:gb_est33 33:gb_est34
34:gb_est35 35:gb_est36 36:gb_est37 37:gb_gss1 38:gb_gss2
39:gb_gss3 40:gb_gss4 41:gb_gss5 42:gb_gss6

Statistics: Mean 10.492; Variance 1.935; scale 5.421

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------|--------|----|----------|------------------------|-----------|
| 1 | 393 | 55.5 | 476 | 18 | AA827795 | od08a05.s1 NCI_CGAP_GC | 0.00e+00 |
| 2 | 390 | 55.1 | 459 | 24 | A1206011 | q914h09.x1 Soares_plac | 0.00e+00 |
| 3 | 379 | 53.5 | 446 | 30 | H02902 | y139f01.s1 Soares_plac | 0.00e+00 |
| 4 | 372 | 52.5 | 469 | 22 | A1095332 | oy17c05.s1 Soares_sene | 0.00e+00 |
| 5 | 356 | 50.3 | 424 | 34 | W47265 | zc39a12.r1 Soares_sene | 0.00e+00 |
| 6 | 309 | 43.6 | 336 | 17 | AA731451 | n294g05.s1 NCI_CGAP_GC | 0.00e+00 |
| 7 | 299 | 42.2 | 309 | 30 | R35439 | yh79d03.s1 Soares_plac | 0.00e+00 |
| 8 | 296 | 41.8 | 332 | 20 | AA856724 | od54c01.s1 NCI_CGAP_GC | 0.00e+00 |
| 9 | 129 | 18.2 | 157 | 17 | AA747293 | nx88a10.s1 NCI_CGAP_GC | 2.48e+00 |
| 10 | 75 | 10.6 | 347 | 10 | AA260340 | va91b02.r1 Soares_mous | 2.51e-99 |

| Result | LOCUS | DEFINITION | ACCESSION | NID | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|--------|--|--|-----------|-------------|---------|----------|--------|----------|-----------|---------|-------|---------|---------|
| 1 | AA827795 | 476 bp mRNA | EST | 25-MAR-1998 | | | | | | | | | |
| 1 | od08a05.s1 | NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1367312 3' | | | | | | | | | | | |
| 1 | similar to TR:064906 | 064906 SIMILAR TO GENBANK ACCESSION NUMBER | | | | | | | | | | | |
| 1 | L26081.. | .. mRNA sequence. | | | | | | | | | | | |
| 1 | AA827795 | | | | | | | | | | | | |
| 1 | 92900158 | | | | | | | | | | | | |
| 1 | AA827795.1 | GI:2900158 | | | | | | | | | | | |
| 1 | EST. | | | | | | | | | | | | |
| 1 | human. | | | | | | | | | | | | |
| 1 | Homo sapiens | | | | | | | | | | | | |
| 1 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | | | | | | | | | | | |
| 1 | Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | | | | | | | | |
| 1 | (bases 1 to 476) | | | | | | | | | | | | |
| 1 | NCI_CGAP | http://www.ncbi.nlm.nih.gov/ncicgap. | | | | | | | | | | | |
| 1 | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | | | | | | | | | | |
| 1 | Tumor Gene Index | | | | | | | | | | | | |
| 1 | Unpublished (1997) | | | | | | | | | | | | |
| 1 | On Jan 14, 1998 this sequence version replaced gi:1877580. | | | | | | | | | | | | |

ALIGNMENTS

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonafido, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bdrp/image/image.html
Insert length: 1737 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 431.
Location/Qualifiers

FEATURES

source

1..476

/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGAAGTGGAGCGCGCCCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="IMAGE:1367312"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 106 a 134 c 147 g 89 t

ORIGIN

Query Match 55.5%; Score 393; DB 18; Length 476;
Best Local Similarity 95.7%; Pred. No. 0.00e+00;

Matches 464; Conservative 0; Mismatches 11; Indels 10; Gaps 10;

Db 1 CCACGAGCCAGGAGCTCTCTGTGTGGAGAGAGCTGCAAGGCTTACTCTTTGA 60
|||||
Qy 117 CCAGGAGCCAGGAGCTCTCTGTGTGGAGAGAGCTGCAAGGCTTACTCTTTGA 176
|||||
Db 61 CTTCCCGGAGGCAAGAACGATCTGTGCGACGGTGAATATGCGCTCCACAAAGGGGTC 120
|||||
Qy 177 CTTCCCGGAGGCAAGAACGATCTGTGCGACGGTGAATATGCGCTCCACAAAGGGGTC 236
|||||
Db 121 CTGCTGATGAAGGGGAGCGGAGAGTACATCATCTCTCTGAGAGGGGAGTGAGG 180
|||||
Qy 237 CTGCTGATGAAGGGGAGCGGAGAGTACATCATCTCTCTGAGAGGGGAGTGAGG 296
|||||
Db 181 GCTGCTGCTGTGGACCAACGCCCGGACCCAGCTGTGTAAGTGGAC 240
|||||
Qy 297 GCTGCTGCTGTGGACCAACGCCCGGACCCAGCTGTGTAAGTGGAC 355
|||||
Db 241 TGTGTGTC-AC-TTGGGAGATGAGAGCTACGCC-TCACGCCGAGAGAAC-TCC 296
|||||
Qy 356 TGTGTGTCACACTGTGGAGAGTACGCCCTTCAGCCCGGAGAGAACGTC 415
|||||
Db 297 C-TGTGCTCTGTT-GAAGGGGAGAGTGTATTCACCATCCGAA-GCA-GGAAT-ACA 351
|||||
Qy 416 CGTGTTCTGTTTGAAGGGGAGCAAGTGTATTCACCATCCGAAAGGATTTACA 475
|||||
Db 352 ATGGGAAGATCCCTCGGTTCCGCGCATCCGGGGGAGAGTGAAGCTTAAACCGATATA 411
|||||
Qy 476 ATGGGAAGATCCCTCGGTTCCGCGCATCCGGGGGAGAGTGAAGCTTAAACCGATATA 535
|||||
Db 412 CTGTGATGAGAAACCAACAGCTTATCAAAAGCCACCATGTGACAAAGACAGGCTTACG 471
|||||
Qy 536 CTGTGATGAGAAACCAACAGCTTATCAAAAGCCACCATGTGACAAAGACAGGCTTACG 595
|||||
Db 472 ATGAC 476
|||||
Qy 596 ATGAC 600
|||||

RESULT

2

LOCUS AI206011 459 bp mRNA EST 29-NOV-1998
DEFINITION gq1409.x1 Soares.Placentae_8to9weeks_2NBHP8to9M Homo sapiens cDNA
clone IMAGE:1759553 3 similar to TR:064906 Q64906 SIMILAR TO
GENBANK ACCESSION NUMBER L26081.; mRNA sequence.
ACCESSION AI206011

NID 93764683
VERSION AI206011.1 GI:3764683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 459)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CCAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2152102.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 2686 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 446.
Location/Qualifiers

FEATURES

source

1..459

/organism="Homo sapiens"
/note="Organ: Placentae; Vector: pT73D (Pharmacia) with a
modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCATCTGAAGTGGAGCGCGCCCTCATTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="IMAGE:1759553"
/clone_lib="Soares.Placentae_8to9weeks_2NBHP8to9M"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/lab_host="DH10B (ampicillin resistant)"

BASE COUNT 87 a 134 c 151 g 87 t

ORIGIN

Query Match 55.1%; Score 390; DB 24; Length 459;
Best Local Similarity 95.8%; Pred. No. 0.00e+00;

Matches 431; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

Db 14 CCGCCGAGGGCCACTAAGAGCGGACCCCGCATCTTCCCGCTGGAAGGCCATGTAG 73
|||||
Qy 8 CCGCCGAGGGCCACTAAGAGCGGATNCTANNCTTCCCGCTGGAAGGCCATGTAG 67
|||||
Db 74 GCGAGACGGGGTGGACCTTTGGCAGACTGAGCGGACACGGTCTTTTCCAGAGCCAG 133
|||||
Qy 68 GCGAGACGGGGTGGACCTTTGGCAGACTGAGCGGACACGGTCTTTTCCAGAGCCAG 127
|||||
Db 134 GCAGCTCCTGCTGTGGGGGAGGAGCGTGGCAAGTCACTCTTTGACTTCCCGAGG 193
|||||
Qy 128 GCAGCTCCTGCTGTGGGGGAGGAGCGTGGCAAGTCACTCTTTGACTTCCCGAGG 187
|||||
Db 194 GCAAGACGATCTGTGCGACGAGTGAATATGCGCTCCACAAAGGGGCTCGTGGATA 253
|||||
Qy 188 GCAAGACGATCTGTGCGACGAGTGAATATGCGCTCCACAAAGGGGCTCGTGGATA 247
|||||
Db 254 AGCGGACTGTGAGAACTACATCACTCTCTGGAGAGCGGAGTGAAGGGCTGCTGCT 313
|||||
Qy 248 AGCGGACTGTGAGAACTACATCACTCTCTGGAGAGCGGAGTGAAGGGCTGCTGCT 307
|||||
Db 314 GTGGCAACAAGCGCGGACCCAGCTGTGGAAGCTTGTGATGGCACTGTGGTCCAC 373
|||||
Qy 308 GTGGCAACAAGCGCGGACCCAGCTGTGGAAGCTTGTGATGGCACTGTGGTCCAC 366
|||||
Db 374 -TTGGGAGATGAGAGGCTATGCCCTTCAGTCCGAGAGAAC-TACCGTGT-TTCTGT 430
|||||

OY 367 CTGGCCGAGTGGAGCTACGCCCCCTTCAGCCCGGACGAGACGTCCCTGTTCT 426

Db 431 TT-GAAGGGGAGAGGTGTATTCACCATC 459
|||||
OY 427 TTTGAGGGGAGAGGTGTATTCACCATC 456

RESULT 3
LOCUS H02902 446 bp mRNA EST 20-JUN-1995
DEFINITION Y139f01.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
IMAGE:151129 3', mRNA sequence.

ACCESSION H02902
NID 9865835
VERSION H02902.1 GI:865835
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 446)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Paterson, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., and
Trevaskis, E., Waterston, R., Williamson, A., Woldmann, P., and
Wilson, R.
The Mashu-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1575
High quality sequence stops: 196
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1575 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 196.
Location/Qualifiers
1. .446
/organism="Homo sapiens"
/note="Organ: placenta; Vector: p773D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AAGTGAAGAATTCGCGCCGACGAGAAATTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo.
/db_xref="GDB:563192"
/db_xref="taxon:9606"
/map="851H08:12:12q24.13-12q24.31"
/clone="IMAGE:151129"
/clone_lib="Soares placenta Nb2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 93 a 115 c 136 g 93 t 9 others
ORIGIN

Query Match 53.5%; Score 379; DB 30; Length 446;
Best Local Similarity 92.9%; Pred. No. 0.00e+00;
Matches 417; Conservative 0; Mismatches 29; Indels 3; Gaps 2;
Db 1 GCCGCTGAAGGCGCATGTAGGAGAGGCGGGTGTGGCCACAGTACGCGCAC 60
OY 46 GCCGCTGAAGGCGCATGTAGGAGAGGCGGGTGTGGCCACAGTACGCGCAC 105

Db 61 ACGGTCTTTTCACGAGCCAGGACGCTCTCTGTGTGGTGGAGAGACGTGGCAAGTTC 120
|||||
OY 106 ACGGTCTTTTCACGAGCCAGGACGCTCTCTGTGTGGTGGAGAGAGGTGCAAGTTC 165
|||||

Db 121 TACCTCTTTGACTTCCCGGAGGCGAAGAACCATGTGTGCCGACCGTGAATATCGGCTCC 180
|||||
OY 166 TACCTCTTTGACTTCCCGGAGGCGAAGAACCATGTGTGCCGACCGTGAATATCGGCTCC 225
|||||

Db 181 ACAAGGGGCTCTGTCTGTATTAAGCGGAGCTGCGAGACTACATCTCTCTGGAGAG 240
|||||
OY 226 ACAAGGGGCTCTGTCTGTATTAAGCGGAGCTGCGAGACTACATCTCTCTGGAGAG 285
|||||

Db 241 CGGANTGAGGGGCTGCTGCTGGCTGTGG-ACCAAGCCCGGCAACCCAGCTGTGAACTGT 299
|||||
OY 286 CGGAGTGAAGGGGCTGCTGCTGGCTGTGGCAACCAAGCCCGGCAACCCAGCTGTGAACTGT 345
|||||

Db 300 GGTGATNGCACATNTGTGCTACCTTGGCGAGATGAGAGCTACGCCCTTTACGCCNAGC 359
|||||
OY 346 GTGAATGCACCTGTGTGCTGACCTTGGCGAGGTGAGAGCTACGCCCTTTACGCCNAGC 405
|||||

Db 360 GAGACTTCC--TGCTCTNTTTTAAGGGGAGCAAGTTATTCANCATCCGNAAGCA 417
|||||

OY 406 GAGACGTCCCGTGTGTCTTTTGAAGGGGAGCAAGTGTATTCACCATCCGNAAGCA 465
|||||

Db 418 AAGATTACAATTGGGAGAGATCTTCGTT 446
|||||

OY 466 AAGATTACAATTGGGAGAGATCTTCGTT 494
|||||

RESULT 4
LOCUS A1095332 469 bp mRNA EST 24-SEP-1998
DEFINITION OY17C05.s1 Soares senescent fibroblasts Nb2SF Homo sapiens cDNA
clone IMAGE:1666088 3' similar to TR:064906 064906 SIMILAR TO
GENBANK ACCESSION NUMBER U26081.; mRNA sequence.

ACCESSION A1095332
NID 93434308
VERSION A1095332.1 GI:3434308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 469)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT On Jan 19, 1998 this sequence version replaced gi:2286378.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1645 Std Error: 0.00
Seq primer: -40m13 fwd, EF from Amersham
High quality sequence stop: 468.
Location/Qualifiers
1. .469
/organism="Homo sapiens"
/note="Vector: p773D (Pharmacia) with a modified
polylinker V_type: phagemid; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTTCACATCTGAAGTGGAGGCGCGCATTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo."
/db_xref="taxon:9606"
/map="13"

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS 1 (bases 1 to 336)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)
JOURNAL On Nov 29, 1993 this sequence version replaced gi:430441.
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbtp/image/image.html
Insert Length: 1045 Std Error: 0.00
Seq primer: -40m13 fwd. E7 from Amersham
High quality sequence stop: 202.
FEATURES
Location/Qualifiers
1..336
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germlinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGAGGCGGCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1303160"
/clone_lib="NCI-CGAP_GCB1"
/tissue_type="germlinal center B cell"
/lab_host="DH10B"
BASE COUNT 66 a 95 c 113 g 62 t
ORIGIN
Query Match 43.6%; Score 309; DB 17; Length 336;
Best Local Similarity 97.3%; Pred. No. 0.00e+00;
Matches 328; Conservative 0; Mismatches 7; Indels 2; Gaps 2;
Db 1 GGGCGAGACCGGCTGGACCTTGGCAGACTGAGCGGACAGCGGCTTTCCAGAGCCA 60
Qy 67 GGGCGAGACCGGCTGGACCTTGGCAGACTGAGCGGACAGCGGCTTTCCAGAGCCA 126
Db 61 GGCACCTCCTCTGTGTGGGTGGAGAGCTGCAAGTCTACTCTTTGACTTCCCGAG 120
Qy 127 GGCACCTCCTCTGTGTGGGTGGAGAGCTGCAAGTCTACTCTTTGACTTCCCGAG 186
Db 121 GGCAGAAACGATCTGTGGCGACGGTGAATATCGGCTCCCAAGAGGGTCTGTGTGAT 180
Qy 187 GGCAGAAACGATCTGTGGCGACGGTGAATATCGGCTCCCAAGAGGGTCTGTGTGAT 246
Db 181 AAGCGGAGCTGGAGAACTACATCTCTCCGAGAGAGCGGAGAGGGGCTGTGGCC 240
Qy 247 AAGCGGAGCTGGAGAACTACATCTCTCCGAGAGAGCGGAGAGGGGCTGTGGCC 306
Db 241 TGTGCGACCAAGCGCGGACCGGACGCTGTGAACCTGTGAATGCACTGTGTGCA 300
|||||

Qy 307 TGTGCGACCAAGCGCGGACCGGACCGGACGCTGTGAACCTGTGAATG-CAGTGTGTGCCA 365
Db 301 C-TTGGCGAGATGAGAGCTACGCCCCCTTCAGCCG 336
Qy 366 CTTGGCGAGATGAGAGCTACGCCCCCTTCAGCCG 402
RESULT 7
LOCUS R33439 309 bp mRNA EST 28-APR-1995
DEFINITION yb79d03.s1 Soares placenta NB2HP Homo sapiens cDNA clone
IMAGE:135941.3, mRNA sequence.
ACCESSION R33439
NID 9789297
VERSION R33439.1 GI:789297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS 1 (bases 1 to 309)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le M., Lennon, G., Marra, M.,
Parsons, D., Rifkin, B., Rohlfing, T., Soares, M., Ten, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The Mashu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1078
High quality sequence stops: 273
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Possible reversed clone: polyT not found
Insert Length: 1078 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 273.
FEATURES
Location/Qualifiers
1..309
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pT73D (Pharmacia) with a
modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'-
AAGTGAAGAAATTCGGCGGCGGACGAGATTTTCTTTTCTTTTCTTTT-3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="GDB:541932"
/db_xref="taxon:9606"
/map="10 q11.2-q22.1"
/clone="IMAGE:135941"
/clone_lib="Soares placenta NB2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 58 a 86 c 104 g 58 t 3 others
ORIGIN
Query Match 42.2%; Score 299; DB 30; Length 309;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 1 CGCGCTGCGGCCGACGAGGCGACCTTAAGAGCGGATNCTTTCGCCGTGCAAGGC 60
Qy 1 CGCGCTGCGGCCGACGAGGCGACCTTAAGAGCGGATNCTTTCGCCGTGCAAGGC 60
|||||

Db 61 CATGTAGGCGACGACGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTCCAC 120
|||||
Qy 61 CATATAGGCGACGACGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTCCAC 120
121 GAGCCAGGACAGCTCTCTGTGTGGGTGGAGAGACGTGCGAAGCTCTACTCTTTGACTTC 180
|||||
Qy 121 GAGCCAGGACAGCTCTCTGTGTGGGTGGAGAGACGTGCGAAGCTCTACTCTTTGACTTC 180
181 CCCAGGCGCAGAACGATCTGTGCGACGGTGAATATCGGCTCCACAAAGGGTCTCT 240
Qy 181 CCCAGGCGCAGAACGATCTGTGCGACGGTGAATATCGGCTCCACAAAGGGTCTCT 240
241 CTGATATAGCGGACGCTGCGAAGTATACATCTCTCTGAGAGAGCGGAGTGAAGGCT 300
|||||
Qy 241 CTGATATAGCGGACGCTGCGAAGTATACATCTCTCTGAGAGAGCGGAGTGAAGGCT 299
Db 301 GCTGGCCTG 309
|||||
Qy 300 GCTGGCCTG 308

RESULT 8
LOCUS AA56724 332 bp mRNA EST 09-MAR-1998
DEFINITION o54601.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1371744 3'
similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
126081. ; mRNA sequence.
AA56724
G2945026
AA56724.1 GI:2945026
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 332)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151919.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 323.
Location/Qualifiers

FEATURES

1.332
/organism="Homo sapiens"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20⁺, IgD⁺),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTACCATCTGAGGTGGAGCGGCGCCATCTTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

/db_xref="taxon:9606"
/map="5: 5023.3-5031.1"
/clone="IMAGE:1371744"
/clone_id="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT 66 a 95 c 109 g 62 t
ORIGIN

Query Match 41.88; Score 296; DB 20; Length 332;
Best Local Similarity 97.68; Pred. No. 0.00e+00;
Matches 324; Conservative 0; Mismatches 4; Indels 4; Gaps 4;

Db 1 GGACTTTGGCGACAGTACGCGACAGCGGTCTTTTCCAGAGCCAGGACGCTCTGT 60
|||||
Qy 81 GGACTTTGGCGACAGTACGCGACAGCGGTCTTTTCCAGAGCCAGGACGCTCTGT 140
141 GTGGGTGGAGAGAGTGGCAAGGTCTACCTCTTTGACTTCCCGAGGCGAAGACCATC 200
Qy 141 GTGGGTGGAGAGAGTGGCAAGGTCTACCTCTTTGACTTCCCGAGGCGAAGACCATC 200
Db 61 GTGGGTGGAGAGAGTGGCAAGGTCTACCTCTTTGACTTCCCGAGGCGAAGACCATC 120
|||||
Qy 141 GTGGGTGGAGAGAGTGGCAAGGTCTACCTCTTTGACTTCCCGAGGCGAAGACCATC 200
Db 121 TGTGCGACGCTGAATATCGGCTCCACAAAGGGTCTGTGATAGCGGACTGCCA 180
|||||
Qy 201 TGTGCGACGCTGAATATCGGCTCCACAAAGGGTCTGTGATAGCGGACTGCCA 260
Db 181 GAATACATCACTCTCTGTGAGAGGCGGAGTGAAGGCGCTGCTGGTGGACCAACGC 240
|||||
Qy 261 GAATACATCACTCTCTGTGAGAGGCGGAGTGAAGGCGCTGCTGGTGGACCAACGC 320
Db 241 ACGGACCCAGCTGTGAGAGTGTGATGAGCTGTGCTGTC-AC-TTGGCAGATGA 298
|||||
Qy 321 CCGGACACCCAGCTGTGAGAGTGTGATGAGCTGTGCTGTC-AC-TTGGCAGATGA 379
Db 299 GAGGCTACGCCC-TTCAAGCCCGGACGAGAAC 329
|||||
Qy 380 GAGGCTACGCCCCTTCAAGCCCGGACGAGAAC 411

RESULT 9
LOCUS AA747293 157 bp mRNA EST 17-MAR-1999
DEFINITION nx88a10.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1269306 3'
similar to TR:064906 064906 SIMILAR TO GENBANK ACCESSION NUMBER
126081. ; mRNA sequence.
AA747293
G2787251
AA747293.1 GI:2787251
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 157)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:693526.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 825 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham

FEATURES High quality sequence stop: 1.
Location/Qualifiers
1..157
source

/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20⁺, 19D⁺),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGAGTGGAGCGCGCTCATTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 33 a 43 c 54 g 26 t 1 others
ORIGIN
/map="1p: 5"
/clone="IMAGE:1269306"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

Query Match 18.2%; Score 129; DB 17; Length 157;
Best Local Similarity 94.6%; Pred. No. 2,48e-213;
Matches 141; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Db 5 GGCAGAACGATCTGTGGCAGCGTGAATTCGCTCCACAAAGGGGTCTGTGAT 64
|||||
Qy 187 GGCAGAACGATCTGTGGCAGCGTGAATTCGCTCCACAAAGGGGTCTGTGAT 246
|||||
Db 65 AAGCGGACTGCGAATCATCATCTCTCTGGAGAGCGCGAGTGAAGGGGCTGCTGC 124
|||||
Qy 247 AAGCGGACTGCGAATCATCATCTCTCTGGAGAGCGCGAGTGAAGGGGCTGCTGC 306
|||||
Db 125 TGTGG-ACCAACGCCCGGACCCAGCTG 152
|||||
Qy 307 TGTGGACCAACGCCCGGACCCAGCTG 335
|||||

RESULT 10
LOCUS AA260340 347 bp mRNA EST 18-MAR-1997
DEFINITION va91002.r1 Soares mouse 3NME12.5 Mus musculus cDNA clone
IMAGE:746763.5 similar to TR:G1000717 G1000717 SIMILAR TO GENBANK
ACCESSION NUMBER L26081.1; mRNA sequence.
AA260340
NID 91896824
VERSION AA260340.1 GI:1896824
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 347)
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Sehellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404539.

CONTACT: Maria M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

FEATURES
source

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:455747
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 296.
Location/Qualifiers
1..347
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: whole fetus; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCATCTGAGTGGAGCGCGCTCATTTTCTTTT-
3'], on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 75 a 102 c 90 g 80 t
ORIGIN
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/tissue_type="fetus"
/clone="unknown"
/sex="unknown"
/clone_lib="Soares mouse 3NME12.5"
/clone="IMAGE:746763"
/db_xref="taxon:10090"

Query Match 10.6%; Score 75; DB 10; Length 347;
Best Local Similarity 91.2%; Pred. No. 2.51e-99;
Matches 83; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Db 1 GAGAGCAACCCCTGACAGACCCGAGCTCTCATGTGTCCCGAGTACCAGT 60
|||||
Qy 620 GAGAGCAATCTGTGACAGATCTGAGGCTCTCATGTGTCCCGTGGCCAGT 679
|||||
Db 61 TGTGAGGGGGGACACAGGTGTGAGAGTTC 91
|||||
Qy 680 TGTGAGGGGGGACACAGGTGTGAGAGTTC 710
|||||

RESULT 11
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPIT cDNA library Oryza sativa
cDNA clone 97SN1787, mRNA sequence.
AA754459
NID 92801165
VERSION AA754459.1 GI:2801165
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.
REFERENCE
AUTHORS 1 (bases 1 to 252)
Nahm, B.H., Kim, J.R., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P.,
Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y.,
Lee, M.C. and Eun, M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797457.

CONTACT: Eun M.Y.
Department of Cyto genetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.astl.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhnahe@bioserver.myongji.ac.kr


```

/db_xref="taxon:4530"
/map="6"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAP11 CDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

```

```

Query Match      7.2%; Score 51; DB 17; Length 247;
Best Local Similarity 13.4%; Pred. No. 3.81e-52;
Matches 33; Conservative 121; Mismatches 87; Indels 6; Gaps 6;

```

```

Db      1  HMDCNNNTWRGCCCCBAMNNKHTHMTBMCVBRVGTNNKNGHRTTWNDCSDNA 60
      CP      686  CTGCACACTGGGCGACCGGACACATTTAG-AGGAGCCTCAGACTTCTTGACGAGATT 628
      Db      61  HCRYTBWYARSKYGYGTBYSMVNDTNGTGKTTVNVHSGMNNRCSNVYVWBT 120
      CP      627  GTCTCTCGAAGAAGTA-CTAGATCTTGCAI-CGTAAGCCTGGTCTTGTCACAGATG 570
      Db      121  AYCDEHYBDRANHVDTFCTNDRCYCNVTASDNG-TSATKRYTGDKTSDCGGCGWRK 179
      CP      569  GTGGCTTTGTAAGTGTGCTGTCATGACAGTACACTGCTGTAACAGCTCACTCTCG 510
      Db      180  VTGSSBIBRCGVNVVRTSMVTDKSTKBSMDMSRVRVHGRMMBNKRGMSRNTD 239
      CP      509  CCC-CGGATCGGGGACCGAGAGATCTCCCAATTGT-NAATCTTGCTTCCGGATGG 452
      Db      240  TKTWTR 246
      CP      451  TGGAATA 445

```

```

RESULT 14
LOCUS      AA754458      247 bp      mRNA      EST      20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAP11 CDNA library Oryza sativa
            CDNA clone 97SN1784, mRNA sequence.
ACCESSION  AA754458
NID        92801164
VERSION    AA754458.1 GI:2801164
KEYWORDS  EST.
SOURCE     Oryza sativa.
ORGANISM  Oryza sativa.
REFERENCE  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
            Poaceae; Oryza.
AUTHORS   Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
            Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
            Lee,M.C., and Eun,M.Y.
            Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
            Unpublished (1998)
            On Jan 14, 1998 this sequence version replaced gi:1797455.

```

```

TITLE
JOURNAL
COMMENT

```

```

Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea, 449-728 bhnamh@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

```

FEATURES

```

1..247
Location/Qualifiers
/organism="Oryza sativa"
/cultivar="Wilyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional CDNA library inserted into lambda ZAP11
vector at 5' end with EcoRI and 3' end with Xho I site."

```

```

/db_xref="taxon:4530"
/map="6"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAP11 CDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT      7 a      16 c      21 g      34 t      169 others
ORIGIN

```

```

Query Match      6.8%; Score 48; DB 17; Length 247;
Best Local Similarity 12.4%; Pred. No. 1.51e-46;
Matches 28; Conservative 103; Mismatches 93; Indels 2; Gaps 2;

```

```

Db      6  MNTVWGCCCCBAMNNKHTHMTBMCVBRVGTNNKNGHRTTWNDCSDNACRYT 65
      QY      21  CCTAAGACGCGGATNCTANNCTTTCGCGCTGGAAGGCATGTAGGCGACGACGGGT 80
      Db      66  VBWYARSKYGYGTBYSMVNDTNGTGKTTVNVHSGMNNRCSNVYVWBTAYCDY 125
      QY      81  GGACTTGGCCAGACT-GAGCCGCACACGGTCTTTTCACAGCAGCAGCAGCTCTCG 139
      Db      126  BHYBDRANHVDTFCTNDRCYCNVTASDNGTATKRVGTGDKTSDCGGCGWRKYTGS 185
      QY      140  TGTGGGTGGAGAGAGCTGCGCAAGTCTACCTCTTTGAC-TTCCCGCAGGCGAAGACGA 198
      Db      186  BYBRGCVNVVVRTSMVTDKSTKBSMDMSRVRVHGRMMBNKRR 231
      QY      199  TGTGCGCAGCGGTGAATTCGGCTCCACAAAGGGCTCTGTCTGG 244

```

```

RESULT 15
LOCUS      R33537      344 bp      mRNA      EST      28-APR-1995
DEFINITION yh79d03.r1 Soares Placenta Nb2HP Homo sapiens CDNA clone
            IMAGE:135941 5' similar to SP:VA93_VACC P21062 PROTEIN ;, mRNA
            sequence.
ACCESSION  R33537
NID        9789395
VERSION    R33537.1 GI:789395
KEYWORDS  EST.
SOURCE     human.
ORGANISM  Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
            The WashU-Merck EST Project
            Unpublished (1995)

```

```

TITLE
JOURNAL
COMMENT

```

```

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1078
High quality sequence stops: 275
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone; similarity on wrong strand
Insert Length: 1078 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 275.
Location/Qualifiers

```

FEATURES

```

1..344
Location/Qualifiers
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pRT3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

```

strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGAGAAATTCGGCGCGAGATTTTATTTTATTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library constructed by Bento Soares and M. Fatima Bonaldo. "

/db_xref="GDB:541932"

/db_xref="taxon:9606"

/map="11"

/clone="IMAGE:135941"

/clone_lib="Soares placenta NB2HP"

/sex="female"

/dev_stage="placenta obtained at birth (full term)"

/lab_host="DH10B (ampicillin resistant)"

75 t 2 others

BASE COUNT 74 a 87 c 106 g

ORIGIN

Query Match

Best Local Similarity 92.3%; Pred. No. 7.18e-34;

Matches 48: Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Db 293 TGAACCTTCCCGACCCCTGTCNCCTTGACACATTGCGCACACGGGACAC 344

Cp 711 TGAACCTTCCCGACCCCTGTCNCCTTGACACATTGCGCACACGGGACAC 661

Search completed: Wed Aug 4 01:29:47 1999
Job time : 1427 secs.


```

/note="sema K1: glycosylphosphatidylinositol-linked
protein; similar to Alcelaphine herpesvirus 1 semaphorin"
/codon_start=1
/product="semaphorin K1"
/protein_id="AAC80456.1"
/db_xref="PID:93930579"
/db_xref="GI:3930579"
/translation="MTPEPPGRAPASPARVPVPPARGLPRLILLMAAASAS
OGHRSGRIFAWKMGHGDDRDGDEGEPPTVLPHEGSSWVGGRKLYLDPER
GKNA SVRTVNI GTKSGCLDRDCENTITLERSSEGLACGINARHPSCKNLVNGTV
VPLEMRGAFPPSPDENSLVLPEDDEVSTLRKEQENKLPERRIRGESELYSDTV
MNPQFIKATIVHDDQADDKYLFEDFENDNADKPEALNARSALCORGOGGEISS
SVSWKMTFLKAMLYCSDPATKRNKFNRIADVLDDPSOMEDTIVGYFSSNVAW
CVSLGIDGDKVFTSLDKGYHSLPNRPPGCLDDPOOIPETFOVADHREVAOR
FMELGLKPLPSHKYHYOKVAVRMOASGHEFVHLYLTDDGTIKHYVEPEQDSHST
PNIEMGIOPFRRAAALIQTMSDAFERKLYVSSOMESVDPDLCVYGGGCGCLMSRD
PYCGMDGRCILSTISSEKSVLOSINPAPEHPECPNFDKRLKLOVSLAPSKRYLS
PMESRHAATYSWNRKENVQSCPEGHQSNQSLITLIENTLAQOYGFCEADQSGYFRE
QHWOQLPEADGIMAHLLGLHACALASMLGVLPLTLLVLYH"

```

| Query Match | Similarity | Score | DB | Length |
|-------------|------------|--|-----|--|
| Best Local | 96.6% | 620 | 31 | 2594 |
| Matches | 681 | Conservative | 0 | Pred. No. 0.00e+00; Mismatches 16; Indels 8; Gaps 8; |
| D | 128 | CCGCCAGGCGCCACCTTAAGAGAGCGGACCCCGCATCTTCGCGCGCTGTGAAAGGCATGTAG | 187 | |
| Q | 8 | CCGCCAGGCGCCACCTTAAGAGAGCGGATINCTANNTCTTCGCGCTGTGAAAGGCATGTAG | 67 | |
| D | 188 | GGCAGAGCCGGGTGACTTTTGGCCAGACTAGACCGCCACACGGGTGCTTTTCCACGAGCCAG | 247 | |
| Q | 68 | GGCAGAGCCGGGTGACTTTTGGCCAGACTAGACCGCCACACGGGTGCTTTTCCACGAGCCAG | 127 | |
| D | 248 | GCAGCTCTCTGTGTGGGTGGAGGACGTGGCAAGTCTACTCTTTTGACTTCCCCGAGG | 307 | |
| Q | 128 | GCAGCTCTCTGTGTGGGTGGAGGACGTGGCAAGTCTACTCTTTTGACTTCCCCGAGG | 187 | |
| D | 308 | GCACAAACGCATCTTGTGGCGACGTGAAATATGGCTCCCAAAAGGGGTCCTGTGAGTA | 367 | |
| Q | 188 | GCACAAACGCATCTTGTGGCGACGTGAAATATGGCTCCCAAAAGGGGTCCTGTGAGTA | 247 | |
| D | 368 | AGCGGACTGCGAGACTACATCACTCTCTGAGAGGCGGAGTAGGGGCTGCTGGCCT | 427 | |
| Q | 248 | AGCGGACTGCGAGACTACATCACTCTCTGAGAGGCGGAGTAGGGGCTGCTGGCCT | 307 | |
| D | 428 | GTGGACCAACAGCCCGGACCCCACTCTGTGAACTGTGTAATGGCACTGTGGTCCAC | 487 | |
| Q | 308 | GTGGACCAACAGCCCGGACCCCACTCTGTGAACTGTGTAATGGCACTGTGGTCCAC | 366 | |
| D | 488 | TTGGCGAGATGAGAGGCTACGCCCGCTTTCACCCCGSAGAGAAC-TGCC-TGGTTCTGT | 544 | |
| Q | 367 | CTTGGCGAGATGAGAGGCTACGCCCGCTTTCACCCCGSAGAGAACGTCCTCGGTCTGT | 426 | |
| D | 545 | TT-GAAGGGAGCAGGTGTATTCACCAATCCGGAA-GCA-GGAAT-ACAATGGAGATC | 600 | |
| Q | 427 | TTTGAAGGGAGCAGGTGTATTCACCAATCCGGAGCAAGAAATTACAATTTGGAGAAAT | 486 | |
| D | 601 | CCTGGTTCGCCGGCATCCGGGGCGAAGTAGTGTGTACACAGATACGTATCTATCGAG | 660 | |
| Q | 487 | CCTGGTTCGCCGGCATCCGGGGCGAAGTAGTGTGTACACAGATACGTATCTATCGAG | 546 | |
| D | 661 | AACCCACAGTTCATCAAAAGCCACATCGAGCCCAAGACACAGGCTTACGATACAAAGATC | 720 | |
| Q | 547 | AACCCACAGTTCATCAAAAGCCACATCGAGCCCAAGACACAGGCTTACGATACAAAGATC | 606 | |
| D | 721 | TACTACTTTCCTCCAGAGAGCAATCCTGTGCAAGAAATCTGAGGCTCTCTCAATGTGTCC | 780 | |
| Q | 607 | TACTACTTTCCTCCAGAGAGCAATCCTGTGCAAGAAATCTGAGGCTCTCTCAATGTGTCC | 666 | |
| D | 781 | CGTGTGGCCCACTTGTGCAGGGGGGACCAAGGTTGGGGAAAGTTCA | 825 | |
| Q | 667 | CGTGTGGCCCACTTGTGCAGGGGGGACCAAGGTTGGGGAAAGTTCA | 711 | |

| RESULT | 2 | AF030698 | 2636 bp | mRNA | PR | 04-SEP-1998 |
|-----------------------|---|--|---------------------|-----------|--------------|-------------|
| LOCUS | | AF030698 | | | | |
| DEFINITION | | Homo sapiens semaphorin L (SEMA1) mRNA, complete cds. | | | | |
| ACCESSION | | AF030698 | | | | |
| NID | | 93523114 | | | | |
| VERSION | | AF030698.1 | GI:3523114 | | | |
| KEYWORDS | | | | | | |
| SOURCE | | human. | | | | |
| ORGANISM | | Homo sapiens | | | | |
| REFERENCE | | Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. | | | | |
| AUTHORS | | 1 (bases 1 to 2636)
Lange, C., Liehr, T., Goen, M., Gebhart, E., Fleckenstein, B. and Enser, A. | | | | |
| TITLE | | New eukaryotic semaphorins with close homology to semaphorins of DNA viruses | | | | |
| JOURNAL | | Genomics 51 (3), 340-350 (1998) | | | | |
| MEDLINE | | 98389619 | | | | |
| REFERENCE | | 2 (bases 1 to 2636) | | | | |
| AUTHORS | | Enser, A., Liehr, T., Lange, C., Goen, M., Gebhart, E. and Fleckenstein, B. | | | | |
| TITLE | | Direct Submission | | | | |
| JOURNAL | | Submitted (21-OCT-1997) Institut fuer Klinische und Molekulare Virologie, Friedrich Alexander Universitaet, Schlossgarten 4, Erlangen, Bavaria 91054, Germany | | | | |
| FEATURES | | Location/Qualifiers | | | | |
| source | | 1..2636 | | | | |
| | | /organism="Homo sapiens" | | | | |
| | | /db_xref="taxon:9606" | | | | |
| | | /map="15p22.3-p23" | | | | |
| | | /chromosome="15" | | | | |
| gene | | 1..2636 | | | | |
| | | /gene="SEMA1" | | | | |
| | | 13..2013 | | | | |
| | | /gene="SEMA1" | | | | |
| | | /note="H-Sema-L" | | | | |
| CDS | | /codon_start=1 | | | | |
| | | /product="semaphorin L" | | | | |
| | | /protein_id="AAC34261.1" | | | | |
| | | /db_xref="PID:93523115" | | | | |
| | | /db_xref="GI:3523115" | | | | |
| | | /translation="MTRPPGGRAAPSPAPARYGPPARLCLPLRLILLMLMAASA
QGHRSGRIFPAWKGHVGDDRDVFGTEPHVLFHPGSSVWVGGRKVVLPDE
GKNASTRVINGSTKSGCDLRDCENTITLLERSBGLAGTNAHPSCMWLVNCTV
VPLDEMGYAPFSPDENSLVLFEDVEYSTIRKQYNGKIPRRRLIGESLYTSDV
MONQEFKATIVHDDQADDKIYFFREDNPDKPEALVNSRAOLCRDQGESL
SVSWNTFLKAMLCVSDAATNNKNNRLOVDYELLPDPGMDRVRVGVSPNNYSV
CVYSLGIDLVFPRSLKGYHSLPNRSPKCLPDOPITPENFOVADRREVAORE
PMQPLTDLPHSKHYKVAVHRKQASHGETHYLYLTITDRGTHIKYVEGEEDHSR
FNMEITPFRRAALITQMSUDAERKRLVSSQWESVDPDLCEVGGCHGLMSLD
PYGWDMDGRICISYSSERSVLSQINSAPREKCPNPKPAKPLQKVSILANSRYVLS
PMBSRAATYSMRKKNENYQSCPEGHQSPNCILFIENTLAQOYHYCEAOEGSFREA
QHMLLDEPDIMEHLIGHACALAAASLMLGLVPLTLGLLVH" | | | | |
| BASE COUNT | | 536 a 835 c 764 g 501 t | | | | |
| ORIGIN | | | | | | |
| Query Match | | 87.6%; | Score 620; | DB 31; | Length 2636; | |
| Best Local Similarity | | 96.6%; | Prod. No. 0.00e+00; | | | |
| Matches 681; | | Conservative 0; | Mismatches 16; | Indels 8; | Gaps 8; | |
| Db 140 | | CCGCCAGGGCCACCTAAGAGCGGACCCGCCATCTTCGCGTGTGGAAAGCCATGTAG | | | | 199 |
| OY 8 | | CCGCCAGGGCCACCTAAGAGCGGACCGATCTNNNTCTTCGCGTGTGGAAAGCCATGTAG | | | | 67 |
| Db 200 | | GGCAGAGACCGGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTCCAGCAGCCAG | | | | 259 |
| OY 68 | | GGCAGAGACCGGGGTGACTTTGGCCAGACTGAGCCGACACGGTCTTTCCAGCAGCCAG | | | | 127 |
| Db 260 | | GCAGCTCTGTGTGTGGGTGGAGAGACGTGGCAAAGTCTACCTCTTTGACTTCCCGAGG | | | | 319 |
| OY 128 | | GCAGCTCTCTGTGTGGGTGGAGAGACGTGGCAAAGTCTACCTCTTTGACTTCCCGAGG | | | | 187 |

| | | | | |
|---|---|--|--|----------------------|
| D | b | 320 | GCAAAAGCATTCTGGCCACGGAATATGGCGTCCCAAAAGGGGTCTGTGGAA | 379 |
| O | y | 188 | GCAAAAGCATCTTGCCACGGTAATATGGCTCCAAAAGGGGTCTGTGGAA | 247 |
| D | b | 380 | AGCGGACTGCAGAATACTACTCTCTTGAGAGGGGAGTAGAGGGCTGTGGCT | 439 |
| O | y | 248 | AGCGGACTGCAGAATACTACTCTCTTGAGAGGGGAGTAGAGGGCTGTGGCT | 307 |
| D | b | 440 | GTGGCACCAAGCCCCGGCACCCCAGTGTGTGAACCTGTGATATGGACTGTGTGCCA | 499 |
| O | y | 308 | GTGGCACCAAGCCCCGGCACCCCAGTGTGTGAACCTGTGATATG-CACTGTGGTCCAC | 366 |
| D | b | 500 | -TTGGCAGATGATGAGGGCGTAAGCCCCCTTCACCCCGGAAGAAAC-TCCC-TGGTTCGT | 556 |
| O | y | 367 | CTTGGCGAAGTGTGAGGCTACGCCCTTTCACCCCGGAAGAAAGTCCCGTGTGTGT | 426 |
| D | b | 557 | TT-GAAGGGAGCAGAGGTGTATTCCACCATCCGGA-ACAAT-ACAATGGGAATC | 612 |
| O | y | 427 | TTTGAAGGGAGCAGAGGTGTATTCCACCATCCGGAAGCAAGAAATTACATTGGGAAT | 486 |
| D | b | 613 | CCTCGGTTCCGCCCATCCGGGGCGAGATAGCTGTACACCATGTACTGTATGCAG | 672 |
| O | y | 487 | CCTCGGTTCCGCCCATCCGGGGCGAGATAGCTGTATACCACCATGTACTGTATGCAG | 546 |
| D | b | 673 | AACCCAGATTATCAAAGCCACATCGTGCACCAAGACACAGCTTAGCATACAAGATC | 732 |
| O | y | 547 | AACCCAGATTATCAAAGCCACATCGTGCACCAAGACACAGCTTAGCATACAAGATC | 606 |
| D | b | 733 | TACTACTTCTTCCGAGAGGACATCCTGTACAAGAAATCCTGAGCTCTCTCATGTGTC | 792 |
| O | y | 607 | TACTACTTCTTCCGAGAGGACATCCTGTACAAGAAATCCTGAGCTCTCTCATGTGTC | 666 |
| D | b | 793 | CGTGTGCCCCAGTTGTGCAGGGGGGACACAGGTTGGGAAAGTTCA | 837 |
| O | y | 667 | CGTGTGCCCCAGTTGTGCAGGGGGGACACAGGTTGGGAAAGTTCA | 711 |
| RESULT | | 3 | | |
| LOCUS | | AF069493 | 2661 bp | mRNA PRI 03-DEC-1998 |
| DEFINITION | | Homo sapiens GPI-anchored membrane protein CDw108 precursor, mRNA, | | |
| complete cds. | | | | |
| ACCESSION | | AF069493 | | |
| NID | | g3947939 | | |
| VERSION | | AF069493.1 | GI:3947939 | |
| KEYWORDS | | | | |
| SOURCE | | human. | | |
| ORGANISM | | Homo sapiens | | |
| Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | | | |
| Primates; Catarrhini; Hominoidea; Homo. | | | | |
| 1 (bases 1 to 2661) | | | | |
| Yanada,A., Kudo,K., Takeshita,T., Harashima,N., Kawano,K., | | | | |
| Sagawa,K., Sugamura,K. and Itoh.K. | | | | |
| Direct Submission | | | | |
| Submitted (02-JUN-1998) Immunology, Kurume University School of | | | | |
| Medicine, Asahi-machi 67, Kurume, Fukuoka 830, Japan | | | | |
| FEATURES | | | | |
| source | | | | |
| location/Qualifiers | | | | |
| 1..2661 | | | | |
| /organism="Homo sapiens" | | | | |
| /db_xref="taxon:9606" | | | | |
| /chromosome="15" | | | | |
| /map="between WI-6247 and CHLC.GAT85D02" | | | | |
| /cell_line="HPB-ALL" | | | | |
| /cell_type="T lymphocyte" | | | | |
| 18..2018 | | | | |
| /note="putative adhesion molecule; lymphocyte activation | | | | |
| antigen" | | | | |
| /codon_start=1 | | | | |
| /product="GPI-anchored membrane protein CDW108 precursor" | | | | |
| /protein_id="AAC82642.1" | | | | |
| /db_xref="PID:g3947940" | | | | |
| /db_xref="GI:3947940" | | | | |
| /translation="MTPTTPGRAPASAPRAVPGPPARLGLPLRLILLILMAAASNA | | | | |
| | | | | |

| Query Match | Best Local Similarity | Matches | 87.6% | Score 620; | DB 31; | Length 2661; |
|-------------|--|---------|-------|------------|--------|--------------|
| | | | | | | |
| Query 8 | CCGCCACAGGCGCCACCTAAGGAGCGGATNCTANNCTTCGCGCGTCTGGAAGGCCATGTAG | 67 | | | | |
| Db 145 | CCGCCACAGGCGCCACCTAAGGAGCGGAGCCGACCTTCGCGCGTCTGGAAGGCCATGTAG | 204 | | | | |
| Query 205 | GGCAGACCGGGGTGGACTTGGCCAGACTAGCCGACAGCGTGCCTTTCCACAGCAG | 264 | | | | |
| Db 68 | GGCAGACCGGGGTGGACTTGGCCAGACTAGCCGACAGCGTGCCTTTCCACAGCAG | 127 | | | | |
| Query 265 | GCACCTCCTCTGTGTGGGTGGAGAGAGTGGCAAGGTCTACTCTTTGACTTCCCGCAG | 324 | | | | |
| Db 128 | GCACCTCCTCTGTGTGGGTGGAGAGAGTGGCAAGGTCTACTCTTTGACTTCCCGCAG | 187 | | | | |
| Query 325 | GCAGAAAGCATCTGTGCGCAGGTGAATATCGGCTCCACAAAGGGTCTGTCTGATA | 384 | | | | |
| Db 188 | GCAAGAAAGCATCTGTGCGCAGGTGAATATCGGCTCCACAAAGGGTCTGTCTGATA | 247 | | | | |
| Query 385 | AGCGGAGCTCGGAAACTACTACTCTCTCTGAGAGAGCGGAGTGAAGGGCTGTGGCCT | 444 | | | | |
| Db 248 | AGCGGAGCTCGGAAACTACTACTCTCTCTGAGAGAGCGGAGTGAAGGGCTGTGGCCT | 307 | | | | |
| Query 445 | GTGGCACCACAGCCCGGACCCAGCTGTGGAACCTGTGTAATGAGCATGTGGTGCAC | 504 | | | | |
| Db 308 | GTGGCACCACAGCCCGGACCCAGCTGTGGAACCTGTGTAATGAGCATGTGGTGCAC | 366 | | | | |
| Query 505 | TTGGCGAGATGAGAGGCTACGCCCTTACAGCCCGGAGCAAGAAC-TCCC-TGGTTCTGT | 561 | | | | |
| Db 367 | CTTGGCGAGATGAGAGGCTACGCCCTTACAGCCCGGAGCAAGAACGTCCGTGGTTCTGT | 426 | | | | |
| Query 562 | TT-CAAGGGGAGCGAGGTGTATTCACCATCCGGA--GCA--GGAT--ACATATGGGAATATC | 617 | | | | |
| Db 427 | TTTCAAGGGGAGCGAGGTGTATTCACCATCCGGAAGCAAGGAATTCATATGGGAAT | 486 | | | | |

| Best Local Similarity | 86.7% | Pred No. | 0.00e+00: |
|-----------------------|--|---|-------------------------------------|
| Matches | 562: | Conservative | 0: Mismatches 77: Indels 9: Gaps 9: |
| Db | 189 | AGGGCAGGACATGTGGACTTTAGCCAGCCTGAGCCACACACCGTCTTTCCATGACC | 248 |
| Qy | 66 | AGGGCAGGACCGGGTGGACTTTGGCCAGACTGAGCCGACACACGGTCTTTCCACAGACC | 125 |
| Db | 249 | GGGCACTCTCTGTGTGGGTGGTGGAGTGGGCAAGGTCTTACCACTTCAACTTCCCCGA | 308 |
| Qy | 126 | AGGCACTCTCTGTGTGGGTGGTGGAGAGCTGGCAAGGTCTTCTTTGACTTCCCCGA | 185 |
| Db | 309 | GGGCAAGAAATGCCCTCTGTGCGCAGGCTGAACATCGGCTCCACAAGGGGTCTGTGAGA | 368 |
| Qy | 186 | GGGCAAGAAAGCATCTGTGCGCACGGTGAATATCGGCTCCACAAGGGGTCTGTGGA | 245 |
| Db | 369 | CAAAAGGAGTGTGGGAATTTACATCTCTTTCTAGAAAGGGGGTAAATGGCTGCTGT | 428 |
| Qy | 246 | TAAAGCGGAGCTGGGAATATCATCTCTCTCGGAGAGCGGAGTGAAGGGGTCTGTGAC | 305 |
| Db | 429 | CTGTGGCACCAATGCCCCGGAAGCCCCAGCTGTGGAACTTGTGAATAGCAGTGGTGAT | 488 |
| Qy | 306 | CTGTGGCACCAATGCCCCGGAAGCCCCAGCTGTGTGAATCTGTGTGAATG-CACGTGGTGC | 364 |
| Db | 489 | GTCACTTGTGTAGATGAAGAGCTATGCCCCCTTCACCCCGGATGAGAAC-TCCT-TGTT | 546 |
| Qy | 365 | A-C-CTTGGCGAGAGTGGAGCTACCCCCCTTCACCCCGGAGAGAAAGTCCCGGT | 422 |
| Db | 547 | CTGTTT-GAAGGAGATGAATGATCTCTACATCCGGAAG-GGA-GGAAT-ACAAACGGAA | 602 |
| Qy | 423 | CTGTTTGAAGGGGAGACAGATGATTTCCACCATCCGGAAGAGGAATTTACAAATGGGA | 482 |
| Db | 603 | GATCCCTGGTTTCGACGCATTCGGGGCGAGAGTGAACATGTACACAAAGTATACAGTCA | 662 |
| Qy | 483 | AGATCCTCGTGGTCCGGCCGATCCGGGGCGAGAGTGAAGTGAAGTACACAGTGAATGTAT | 542 |
| Db | 663 | GGAGAAACCCACAGTTTCATCAAGGCCACCATTTGTCAACCAAGCAAGCTTATGATGATA | 722 |
| Qy | 543 | GGAGAAACCCACAGTTTCATCAAGGCCACCATTCGTGACCAAGCAAGCTTATGATGACAA | 602 |
| Db | 723 | GATCTACTACTTCTTCCGAGAGCAACCCCTGACAAGAACCCCGAGGGTCTCTCAATGT | 782 |
| Qy | 603 | GATCTACTACTTCTTCCGAGAGCAACCATTCGTGACCAAGCAAGCTTATGATGATGATA | 662 |
| Db | 783 | GTCCGAGTAGCCCAAGTTGTGCGAGGGGGGACAGGGGTGTGAGAGTTC | 830 |
| Qy | 663 | GTCCGAGTAGCCCAAGTTGTGCGAGGGGGGACAGGGGTGTGAGAGTTC | 710 |
| RESULT | 5 | | |
| LOCUS | AB017532 | 2608 bp | 09-MAR-1999 |
| DEFINITION | Mus musculus mRNA for msemk1p, complete cds. | | |
| VERSION | AB017532.1 | GI:4519587 | |
| KEYWORDS | msemk1p. | | |
| SOURCE | Mus musculus postnatal brain cDNA to mRNA. | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; | | |
| AUTHORS | Eutheria; Rodentia; Sciurognathia; Muridae; Mus. | | |
| TITLE | Sato, Y. and Takahashi, H. | | |
| JOURNAL | Molecular cloning and expression of novel Semaphorin family of | | |
| REFERENCE | Unpublished (1998) | | |
| AUTHORS | 2 (bases 1 to 2608) | | |
| TITLE | Takahashi, H. and Sato, Y. | | |
| JOURNAL | Submitted (08-SEP-1998) to the DDBJ/EMBL/genbank databases. Hiroshih | | |
| FEATURES | Neurobiology Lab; II Minamioyga, Machida, Tokyo 194, Japan | | |
| | (E-mail: hiroshih@labra.1s.m-kagaku.co.jp, Tel: 81-427-24-6211, | | |
| | Fax: 81-427-24-6314) | | |
| | Location/Qualifiers | | |

```
source 1. .2608
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /dev_stage="postnatal"
        /tissue_type="brain"
gene 6. .2000
       /gene="msemk1"
       /gene="msemk1"
       /gene="msemk1"
       /product="msemk1p"
       /protein_id="BA05665.1"
       /db_xref="PID:d1039405"
       /db_xref="PID:q4519588"
       /db_xref="GI:4519588"
CDS 6. .2000
     /codon_start=1
     /translation="MPPPPGARAAPSAPRRLSLPARGCLRLRLLLVFWAASA
OCHSRGPRISAWKQDHDVDSQEPHVLTHPEPSESVMWGRKVIHNPPEGN
ASRVNIGSTKSCODKODCGNYITLLRRKNGLLVCTNARKPCNMLVNDVYMS
LGEMGYAFPSDENSLVLEGEDEVSTIRKQENKILPRFRIGESGLSDTVMO
NPOFIKATIVHODQADKLYFFREDNDKNPEAPLANSRAVQLCRGOGESSLSV
SKMNTFLKMLVCSDAATNRNRLQDVLLDPDSQMDRTVGVFSPNMYSAVCV
YSLGIDIRFERTSLKGYHMGJPNRPNGCLPKKQPIPETOVADSHFVQARPEM
GPLKTPLEHSHKHQVYVHRQASNGETFLVLLTDDGTHKAVESDDHSEVEN
IMEIQPFHRAAIAISLDADRKLVTQWVSOVPLDMCEVSGGCGCLMSDIPY
CGMDQRCVSIYSQSRVQSINPAEPHRECNPPDEAPLQKVSILARNSRYLLCPM
ESRHATYLRHEENVEOSCEPQHSPLCIENLTARQYGHYRCBAQDSYLRFAQH
WELPEDRALAOLMGHARALAAFMVLGVPTLLIGLVH"
BASE COUNT 554 a 769 c 727 g 558 t
ORIGIN
Query Match 60.9% Score 431; DB 32; Length 2608;
Best Local Similarity 86.7% Pred. No. 0.00e+00;
Matches 562; Conservative 0; Mismatches 77; Indels 9; Gaps 9;

Db 182 AGGCGAGACATGTGACTTTAGCCAGCTGAGCCACACACCGCTTTTCATGACC 241
    |||||
    66 AGGCGAGACCGGGTGAAGCTTTGGCCAGACTGACCGACCGCTCTTTCCACGAGCC 125
Db 242 GGGCAGCTTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 301
    |||||
    126 AGGCGAGCTCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
Db 302 GGGCAGAAATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 361
    |||||
    186 GGGCAGAAAGCATCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 245
Db 362 CAAACAGACTGTGGAATTCATCACTCTTGTAGAAAGCGGGGTAAATGGCTGTGGT 421
    |||||
    246 TAAGCGGAGCTGGAGAACTACATCACTCTCTGAGAGGCGGAGTGGAGGCTGCTGGC 305
Db 422 CTGTGACCAATGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 481
    |||||
    306 CTGTGACCAATGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 364
Db 482 GTACATGTGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT 539
    |||||
    365 A-C-CTTGGCGAGAGTGGAGGCTACGCCCTTACGCCCGGAGAGAGAGTCCCGTGTT 422
Db 540 CTGTTT-GAAGAGATGAAGTGTACTTACATCCGAA-GCA-GGAAT-ACAACGGGAA 595
    |||||
    423 CTGTTTGAAGGGAGGAGAGTGTATTCACATCCGAAAGCAAGGAATTAATATGGGA 482
Db 596 GATCCCTCGGTTTCAGAGCATTTGGGGCGAGAGTGAAGTGTACACAGATGATACACTCAT 655
    |||||
    483 AGATCTCGGTTTCAGAGCATTCGGGGCGAGAGTGAAGTGTATACACAGATGATCTCTCA 542
Db 656 GCAGAACCAAGTTCATCAAGGCGGAGGAGTGTGACACCAAGGAGGAGGAGGAGGAGGAGG 715
    |||||
    543 GCAGAACCAAGTTCATCAAGGCGGAGGAGTGTGACACCAAGGAGGAGGAGGAGGAGGAGG 602
Db 716 GATCTACTACTTCTTCGAGAGAGCAACCTGACAGAGAACCCGAGGCTCTCTCATGT 775
    |||||
    603 GATCTACTACTTCTTCGAGAGAGCAATCTGACAGAGAAATCTGAGGCTCTCTCATGT 662
```

```
Db 776 GTCCCGAGTAGCCAGTGTGACGAGGGGAGCCAGGAGTGTGAGAGTTC 823
    |||||
    663 GTCCCGTGTGGCCACTTGTGACGAGGGGAGCCAGGAGTGTGAGAGTTC 710
FEATURES
source 1. 8861
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /map="15p22.3-p23"
        /chromosome="15"
        join<1. .114.605. .646.944. .1036.1238. .1322.1443. .1553.
        2176. .2315.2907. .3091.3945. .4053.4146. .4344.6876. .7003.
        7178. .7332.7533. .7594.7904. .>8265)
        /gene="SEMA1"
        /product="semaphorin L"
        /evidence="experimental"
        <1. .>8265
        /gene="SEMA1"
        join<1. .114.605. .646.944. .1036.1238. .1322.1443. .1553.
        2176. .2315.2907. .3091.3945. .4053.4146. .4344.6876. .7003.
        7178. .7332.7533. .7594.7904. .8265)
        /note="sequence starts with the putative second exon of
        SEMA1; H-SEMA-L"
        /codon_start=1
        /product="semaphorin L"
        /protein_id="AAC34741.1"
        /db_xref="PID:q351779"
        /db_xref="GI:351779"
        /translation="EPHVLTHPEPSSSSVMWGRKVIHNPPEKNASVRVINGST
KGSCLDRKDCENVITLLERKSEGLACITNARHPCNMLVNGTVPLDEMGTAFSP
DENSLVLEGEDEVSTIRKQENKILPRFRIRGSESLYSDTVQONPOFIKATIVHQ
DOAYDRIYFFREDNDKNPEAPLANSRAVQLCRGOGESSLSVSKMNTFLKMLV
CSDAATNKNFNRLODVFLLPDSQMDRTVGVFSPNMYSAVCVSLGIDIRFERT
SLKGYHSHLNPRNGCLPKKQPIPETOVADSHFVQARPEMGPLKTPLEHSHKHQV
HYOKAVYHRMQASHETFLVLLTDDGTHKAVESDDHSEVENIMEIQPFHRAAIAIS
IOTSLDARERKLVSQWESQVPLDCEVYGGGCHCLMSRDYCGMDGRCTISY
SSERSVLDSINPAEPHRECNPPDEAPLQKVSILARNSRYLLCPMESRHATYLRH
ENVEOSCEPQHSPLCIENLTARQYGHYRCBAQDSYLRFAQHOMLLPEDGIMAE
HLHGACLAASLWLVGVPTLLIGLVH"
BASE COUNT 1749 a 2358 c 2757 g 1997 t
ORIGIN
Query Match 20.2% Score 143; DB 31; Length 8861;
Best Local Similarity 97.4% Pred. No. 1.14e-102;
Matches 147; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

| DB | 2176 | ACCCACAGTTCATCAACAGCCACCATGCTGTGACCAAGACCGGCTTACGATGACAAGATCT | 2235 |
|----|------|---|------|
| QY | 548 | ACCCACAGTTCATCAACAGCCACCATGCTGTGACCAAGACCGGCTTACGATGACAAGATCT | 607 |
| Db | 2236 | ACTACTTCTCCAGAGAGCAATCTGCACAGAAGATCCTGAGGCTCCCTCAATGTGCC | 2295 |
| QY | 608 | ACTACTTCTCCAGAGAGCAATCTGCACAGAAGATCCTGAGGCTCCCTCAATGTGCC | 667 |
| Db | 2296 | GTTGCGCCAGTTGTCAGGCGTGAACACGGG | 2326 |
| QY | 668 | GTTGCGCCAGTTGTCAGGCGGAGACACGAGG | 698 |

[illegible]

Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae.

REFERENCE 1 (bases 1 to 130608)

AUTHORS Essner, A., Pflanz, R. and Fleckenstein, B.

TITLE Primary structure of the alcelaphine herpesvirus 1' genome

JOURNAL J. Virol. 71 (9), 6517-6525 (1997)

MEDLINE 97404659

REFERENCE 2 (bases 1 to 130608)

AUTHORS Essner, A., Pflanz, R. and Fleckenstein, B.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-1997) Institut fuer Klinische und Molekulare Virologie, Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Schlossgarten 4, Erlangen 91054, Germany

LOCATION/Qualifiers

SOURCE 1. 130608

1. 130608

/organism="Alcelaphine herpesvirus 1"

/strain="C500"

/specific_host="Connochaetes taurinus taurinus"

/db_xref="taxon:35252"

/lab_host="MDBK bovine epithelial kidney cells; ATCC CCL-22"

918. 1211

/note="putative ORF"

/codon_start=1

/evidence=not_experimental

/product="A1"

/protein_id="AAC58052.1"

/db_xref="PID:g2337968"

/db_xref="GI:2337968"

/translation="MRLHFLFYRSIEFYMKSKHKIFCITLSLYFLTAHKRRARPITDFSRIMGCPCEFKMSRHOLLVNLQMLSPCRNKCSTLSADOCITANEDTYS"

1248. 1254

/evidence=not_experimental

complement(1300..1305)

complement(join(1796..2230,2213..2477))

/note="spliced cDNA detected by 5' RACE-PCR"

/codon_start=1

/evidence=experimental

/product="A2"

/protein_id="AAC58053.1"

/db_xref="PID:g2337969"

/db_xref="GI:2337969"

/translation="MSQNSNSNPSPKRRKYKMKDTEOEKERRSINRRASKNFKRRRIEEOEKGLINLKENSRLRCOVERKRDRLRLRWLNHKKITLQNYNGPPEPRKVENSLKMCATATFLNDQOYTNNLNLPKVSQNTNGFAATATLHNCYKTLANNINFEAKLNCCEVLPSTSLDLDLSIDWNNTLYL"

3492. 5453

/note="A3: AHV-sema, similar to Vaccinia A39"

/codon_start=1

/evidence=not_experimental

/product="semaforin homolog"

/protein_id="AAC58054.1"

/db_xref="PID:g2337970"

/db_xref="GI:2337970"

/translation="MAYLNATVSPVLSLSKVKLFEEGGEGGGLITTEVYHPAMGTCVSRILMLISATTAASKSPIDKPRILVNTDGGQHRGHPGDEPRLFHSLSMSDYVVGNNNTIYLFDFAHSSNASTALINITSNTNRLSTGNCFTLLHNOJDGLACGTNSQKPSQWLNINLTTOFLGKGLAPSPSSGALVLEFDODYSTINLYKSLGSHKRRRIRAGVOELTSDTAMHRPOVQATAVKHNSYDRIYFFQNSHSDFOFPHVPRVGQVSSDQGESLSYKWTTLKARLACVVDGTGRIVYELQDIFIMQAPENSEMETLIYGLSPMNFSAVCETWKIDHVEKTSKLNKYNHKLPTPRGCMKHHOHVTEFEQVADRYPEVADPVQKKNAMPRIIOSKYTYTKLIYRVYVGFPNATIFYLTTIKGTIHLIYVREDSNTALILEINFPQKRAPQIQLNDNTNKLKLVSEMEYS EVDLDGCVGNDGSCFEMSRLPCTWYNTNCSFKQSVETGPGPARILTSEMGDHYAPTIVKHQVSIPLISNYSLSQPAVSNADYFWKDGFTKRCRCHVKHKNDICILLIANS TTATNGTHVCMKEDSVTKLLEVNVTIM"

5590. 5597

5732. 6097

/note="ORF containing putative signal peptide; detected by 5' RACE-PCR"

/codon_start=1

/evidence=experimental

/product="A4"

polya_signal

CDS

/protein_id="AAC58055.1"

/db_xref="PID:g2337971"

/db_xref="GI:2337971"

/translation="MVAOLYHNLITLIASIIYIFFNAPATLYEDDEDFNSREYDF LBSDSQFPRRNSCNCCTREARTRITLLEALKRALQIIAGIINETDIELPPTPIPTMT PLETTPLDTSPPVPSAIP"

6123. 6128

/evidence=not_experimental

6333. 10442

/note="ORF03: similar to H. salmieri and EHV2 ORF3, similar to ORF75"

/codon_start=1

/evidence=not_experimental

/product="tegument protein"

/protein_id="AAC58056.1"

/db_xref="PID:g2337972"

/db_xref="GI:2337972"

/translation="MTRRELQLAVTAEPSSALETAIAALRSIPGLONVITOTADVL VPFSTSPRLREHQLKTELFLFKALRTFDEQOVLPIPLASLSSFTFYGPDIRNL PTQSNMELAILOTBDAKEVENVRIHCRLLFFSGSGEOLNTHVYALREILCGD LTIYQPRNLEFNSLDIFYATIPPRFELDKVRYHYHGDININITGAASYMNASPHT AYTDSEIFRPIPTHVYRFIDRMEIRVRVDPVGHSGRVFRHOGCAIYICSTLNTT GATGCLQKGNTISFISQPTLGQLITTEPWCQGGQELSATATQMLTKTANVETAFK TLMYAGVPLVOGFTLSPLTRERKMAFTGSLITGLPYLILPTAETLIDAMERNR OLLIVEGYPDYEGKRLHNPVNINMSSESRHAHILQALROLMIITPFQVHICCDM QYNEKVTMRKIAKACGDLGMSARSILPHTTLRLKMNERNLNYHRIIKNSMLKVD SAAIIVIVADNEYODYSKRMIDALSGMCGPGLTGLTTPNSNTIYISDRNOYEIYDI QYKMHMPKOPSEGETDTP LAPDMSNTQLKNDVLTDLQIYLRHPTVCKAHITAHV DRCNGHIAQOPGVGPFDPICDFSVYVNLVNDGDI REGESVPRVADWRVARRLI EITYSTPGDITDATALANNIGYHLIPSEHQVYSEKRCNGCIGIEKTFETORDPL GTILAVSECTNCLIGPVENYEELGLSISYEGEIHREYVNSIIMAKKDFCSNMF GVOVNSAENGNCILRSVATAANAPCVPGPSLKPYPKQGSATLRVNLHREHLSGI CMAASGISASETPTPTPSOLRNLLOEMLVYKNAIENLASHGVSDGLICANGCMFAG GLSARLIHDEDEEPYFPIRSETPGVLEYNADVAITARANLYVECOIEYVES DITTVHQNTQLSVLVSRLKHNMTLESVDLLYKEDQVLEETSYGVEHVLVD PYSLIISQSTTRPVLVALLPGCGYDPAALALNNSCFSDTVVPCCKYKHNHREDA PGGPILADDFIAGIYGLSSNIDSDVGDSTIRQMLANNROYINVRNLAKAGSFLTAI GOLACRILFATKAIIGFDAGSOQTFPLIPNARSRYESRWLNEFKIPEDTKAAVFDLRC VLPKWOVGTGLGFSNHNITPEGDLFTROOVAFTPVOSGPARVETPPEAEHPY AGCSESDGRLHALPDPCLAFNTWQOHQNTGCGSELYSPKMLFRLYNSKSKHQ HRSRLRTNLRLHFTENETQQLDPLHADRGAVPDHPQYIPMDAMPDSQ"

11134. 11139

/evidence=not_experimental

11260. 14643

/note="ORF06"

/codon_start=1

/evidence=not_experimental

/product="major ss DNA binding protein"

/protein_id="AAC58057.1"

/db_xref="PID:g2337973"

/db_xref="GI:2337973"

/translation="MALKHQLNHAVEDNIGSKAPIGPCGFIIYPTTHNFKEISLIG DKYAEAGASFLPLHGVTEAEVFPVNVKVAKKIDITVTSVKISTYNYNAIIFHNNEK FESIFSGPGLSGICEKACDLFGVYPTPDEGSETDSCLPWQOKDAMAVITGEG FKRLHLGLTIYKSSQMSVMIKKTEVYRPLTDEDLFIKKSRLRLYIPAVEEYLY TLTYSIOLSRVHNNAIYEAIOECPHVKYKAKIYSREKVALVAGCDTITMIDF AYAAEIGLSYSLSFEAPROBKTVQDYIYQMDITFASCESTDREALKSKMALHIIH AQLEFNSIYVVARVAQPIPSKQVPEVYNVSYIYQHGIANCEETLEDGSPATG APASLSLDGSGFTLOHLYAAASPNLAKRCYILOGCOHQKSTLNAYINTEVFGAA NSVSCALGSGQCPCVCINTLYRLKQRPVLOGSRDYPVIGIINVNEDLDFLFC ASPRDEDEONOTEEPTRYTWOLONTLEKLEAALVSPVADGAGSGSNLLEKF VRRFSDILVLAEAKPTINTMIKNVNFKESIKGSHYIOVNCNINWYMAPCSIMNL YRSILITTIIDIALPISTYSESNPQCGKPNEMLHLHQTMLNFKSFIDKGVITG TEKKVYHAQPSDFEVDADATNNMSPVAVQVRLPAQVLALEKNIKVKNRILFSGP SEHYQNAFLKTNBRNDNYILAGYVFLRSPHOLEPNAKISCLYLMASNCKKQJPC VPVSAEALNKFEFSYINNNSKOFEEVYNMADVDSYVYAKOGLNNAIIRACQOTFY AVTIIHSIFPKVOETCALEYPHVIGTSVSDVDEYVANNONLAKALYINSSIRSAANLA RSRPIVTLIPVVNKKYGIAGNALQOSALQYFGMVGSGVNDKILGSLFPKKOONSMR KRYLEMTPIYGNLKRSTYHOGTAIEIVYKRTIQLSLEDQADEVLNRYVCEIYKSL GAGCADTLDDIOTIIGYSTGMSFSENTERLKDLOLRGTPWTHEMASVLSKSGCETDE YQVFVAEEEOVKLTSMDHSGKVGKRRIAITMFDLDL"

14725. 16767

/note="ORF07"

polya_signal

CDS


```

/codon_start=1
/evidence=not_experimental
/product="transport protein"
/protein_id="AAC58058.1"
/db_xref="GI:2337974"
/db_xref="GI:2337974"
/translation="MGQYLALYLSQIYGLCLDYLVEFCCKPTSLCTRADIACNKKVH
IHEVVASLLOONSLEACALSELSHLENLKTLFCFYIHALNDPNFYFSKHSVLG
CDLKRKNLQVQFNECGIEVNLTLINDLERLSKTPYFVYALNDPNFYFSKHSVLG
QLRISFVPRDIITYITSSCECLEVETSVVPMOGETINELLNLNCHHLVERPPEPT
KGLFESLQNLGLKVLHATDITIEQSVKHEAVLOESLALYKAHTFNNTPOVELISL
LWNSQNPQSDGVCSELKISWSENELQKRPKLNNPEPGHFDLPSPGTELD
LFCGSISSHTDITTLAKQDCSNFMKOTRLTGAKNNPEFMLSNIIXGEVTPRL
KQESAKITCDQSDASKNOYLQAEELKRAYLNLSKEGRKLOACI.STHEEMNSOL
SLKIKGSVYIKOSATILNLHFLFROSWTQASLPVSGSPQEFNSKFTISSLYKSL
SRELSTRLHNFHALITGPTLTOGELFSPENVOALCELAHMHPOKMLNEMIRP
TMEQDQICSNFNEFEYIHERTDLNGVQECMKIRELVYALYNIEMENKLCYTRD
HSCPTACSGSIGKIEYVYTESHAPLILVYSKKWIFKDYALYALYAHQNLNNGNR"
15048..15053
polyA_signal
/evidence=not_experimental
16751..19315
/note="orf08; similar to EBV BALF4, CMV UL55, HSV UL27"
/codon_start=1
/evidence=not_experimental
/product="glycoprotein B"
/protein_id="AAC58059.1"
/db_xref="PID:g2337975"
/db_xref="GI:2337975"
/translation="MAHTGSHVCAFLIFAVLKNVFCOTPTSSSEVEDIVEANTVSDNN
IIRORNNKAGIHSDFSAFPRVCSASNIGDIRFOTSHSCPRTKQKHEHGILLF
KENYVPFVKRYKRIIVTSTIYNGLYADAVTNQHFVSKSPYIETRMDTIYOCYN
SLYTVGGLNLYRTDNDGNSMTVDLOVDGLSNRYRHSHQPEIHAHPGMLIGYRRR
TYNCEVETEDRAAVPPFRFTINIGTIESSPVSAMNEDESSEDPRLIYAKDY
RVYDKFERGTQPOGHTRIEVDKEEYTLISMAQFNIISYCRAMKSDNAIKEHGKS
LHFANDITSAFYEPNTQREVLEKHAVCNLTIESELKSRALKNDHDSNGTQOYLL
TNGGLLWQALPVOOKLIDAKGLDAKCOONTYTTTTRSRORRSVSGIDVYAT
AESITLITQIOPAVDTRAOINNLVLSRAMCGEHRASLIMNELSKINPTSVSSSI
YGRVSKARKIGDIVSYSHCYVVDSDSLSRMSMRPQDRTHEYSSRPYFKTINDS
HLKGLGVNNHILITTAVEICHENLEHTFQGGNNMTFTKNRHVTMPVGDATLTD
TFWNLTLJLVENIDFOYIELYSREKEMSTAFDLETFRENNYTORVYGRRLTDL
ATPNNQFVDAFENGLMDLGVGKTVLNAVSVALFESSIGLINFKNFGGMLLFG
LIAAVTVTILLNKAKRFAQNPVQMIYDPIKTITTSORELOVPIKSHELDRLMLAM
HDVHASKQESKODEBEGSTTSGPADMLNKAKNVLRRRAGKPKLRIDSESTYCP"
19314..19319
polyA_signal
/evidence=not_experimental
19428..22508
Note: remainder of annotations omitted.

Query Match          6.5%; Score 46; DB 37; Length 130608;
Best Local Similarity 61.9%; Pred. No. 1,608-16;
Matches 120; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Db 4188 GAATATACAGAGTGAACACCGCCATGACACGGGCACAGTTTGTCCAGGACACAGCTGTG 4247
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 GAGCTGTACACCAAGTACGTCTATGGAAGAACCCACAGTTTCATCAAGCCACCATCTGTG 576
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4248 CATAAATATGATCTTATGATGATAAATCTACTTTTCTCTTCAAGAAACAGGCACAGT 4307
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 577 CACCAAGACCAAGGCTTACGATGACACAAGATCTACTTCTTCCGAGAGACAAATCTCTAC 636
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4308 GACTTCAAAAGATTCCACATACGTGTACCTAGAGTGGGGGACAGGTGCTCTAGTGATCA 4367
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 AAGATCTCTGAGGCTCTCTCAATGTGTCCGTGTGGCCACAGTTGTGACGAGGGGACACAG 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4368 GGTGGGAGAGACTC 4381
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 GGTGGGGAAGTGT 710
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

[illegible]

CP 367 GGTGCACACAGTCATTCACAGGTTCCAGAGCTGGGGTGGCGGCTGTGTCAC 308
Db 1244 YY 1303
CP 307 AGGCACGACGACCCCTCTACTCCGCTCTCCAGAGAGTATAGTTCTCCGAGTCCGCT 248
Db 1304 YY 1363
CP 247 TATCCAGACAGACCCCTTTGTGGAGCCGATATTCACCGTCGCACAGATCGTCTTGC 188
Db 1364 YYY 1367
CP 187 CCTC 184

RESULT 11
LOCUS MWU87256 1056 bp DNA MAM 02-JAN-1999
DEFINITION Mustela vison GI dinucleotide repeat, chromosome 1q.
ACCESSION U87256
NID 94099442
VERSION U87256.1 GI:4099442
KEYWORDS
SOURCE American mink.
ORGANISM Mustela vison
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Carnivora; Fissipedia; Mustelidae; Mustela.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Brusaard, K., Shukri, N.M., Malchenko, S., Koroleva, I. and Lohi, O.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1997) Breeding and Genetics, Danish Institute of
Animal Science, Blichersalle K25, Tjele 8830, DK
FEATURES
source
1. 1056
/organism="Mustela vison"
/db_xref="taxon:9667"
/chromosome="1"
/map="1q"
/note="primers: 1167F: agccctgcatactactctt, 1167R:
gagatctaccgcgtgttag"
primer_bind
98..119
/standard_name="1167F"
/complement(300..320)
primer_bind
/standard_name="1167R"
BASE COUNT 211 a 221 c 210 g 225 t 189 others
ORIGIN
Query Match 4.9%, Score 35; DB 23; Length 1056;
Best Local Similarity 19.7%; Pred. No. 3.64e-08;
Matches 30; Conservative 66; Mismatches 55; Indels 1; Gaps 1;

Db 559 GMSCKMTMTSGCCSCMRDVBMSKMKWCKSKCKSTGCKDKMSGCATCGKMRCSR 618
QY 292 GAGGGGCTGCTGCTGTGGACCAAGCCCGACCCCTCTGGAACCTGGAAT 351
Db 619 YMMARMVSGSTWARMCCCKSKSGDKSHCKSKRYKMDRYBCKSMCAMVGVKWSAG 678
QY 352 GCACTGTGTGGCCACCTTGGGAGAGTGGAGGCTA-CGCCCTTCAGCCCGAGAGAA 410
Db 679 TGCMAKGMRTWSSCTSTKGAARRKGMARRK 710
QY 411 CGTCCGCTGTCTGTGTTGMAAGGAGAGAA 442

RESULT 12
LOCUS E04076 565 bp RNA PAT 29-SEP-1997
DEFINITION gDNA encoding envelope region of type C hepatitis virus.
ACCESSION E04076
NID 92172286
VERSION E04076.1 GI:2172286
KEYWORDS JP 1992349885-A/1.
SOURCE Hepatitis C virus.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

REFERENCE 1 (bases 1 to 565)
AUTHORS Moriaga, T., Chayama, K., Kumada, H. and Ichikawa, Y.
TITLE NUCLEIC ACID FRAGMENT WITHIN ENVELOPE REGION OF HEPATITIS C VIRUS
AND METHOD FOR DETECTING THE SAME
JOURNAL Patent: JP 1992349885-A 1 04-DEC-1992;
TEIJIN LTD
COMMENT OS Hepatitis C virus
PN JP 1992349885-A/1
PD 04-DEC-1992
PF 29-MAY-1991 JP 1991152169
PI MORINAGA TSUTAE, CHAYAMA KAZUAKI, KUMADA HIROMITSU, PI
ICHIRAWA YATARO
PC C12N15/10, C12O1/68, C12O1/70//C12N15/11;
CC strandedness: Single;
CC topology: Linear;
FEATURES
source
1. 565
/organism="Hepatitis C virus"
/db_xref="taxon:11103"
BASE COUNT 60 a 93 c 107 g 85 t 220 others
ORIGIN
Query Match 4.8%, Score 34; DB 25; Length 565;
Best Local Similarity 27.5%; Pred. No. 1.91e-07;
Matches 44; Conservative 54; Mismatches 61; Indels 1; Gaps 1;

Db 227 CBGGGTGTRBCTGTGTCGSGARRRYNNYTBMGNTGTGGGTCCTBACYC 286
QY 255 CTGGGAGACATACATCTCTCCGAGAGCGAGAGGGCTGTGCTGTGGCAC 314
Db 287 CCAGSTYGCVCRCAGRAVBAVYVYVYCCRCBRYGVRMTGCBGCAVTCGAY 346
QY 315 CAACGCCCGCACCCACGCTGCGAA-CCTGGATGCACTGTGTGCCACTTGGCG 373
Db 347 TGCTYTHGGGSRVGYRCYTCGTGTCGSCYVYTAVGT 386
QY 374 AGAGTGAAGGCTAGCCCTCTAGCCCGAGAGAACT 413

RESULT 13
LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID 93977523
VERSION AR024229.1 GI:3977523
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace, F. Paul, Harris, W.J., Carr, F.J., Old, L.J., Welt, S. and
Kitamura, K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES
source
1. 965
/organism="unknown"
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN
Query Match 4.8%, Score 34; DB 25; Length 965;
Best Local Similarity 14.9%; Pred. No. 1.91e-07;
Matches 26; Conservative 80; Mismatches 66; Indels 2; Gaps 2;

Db 783 RHTVHSGVYSTCTASPYTTSYG-WYRGWGDYGGGTYNNYNGRVTMADTSSN 841
QY 226 ACAAGGGCTCTGTGATTAAGCGAGACTGCGAGACATCACTCTCCGAGAGG 285
Db 842 SRSSVTADTAIVYCYRGARVSDGDDYMGTTVYSSHTVKMTSSSSASVDDRYTTCR 901
QY 286 CGAGGTGAGGGCTGTGCTGTGG-CACCAAGGCCCGCACCGACGTGTGGAACCT 344

This Page Blank (uspto)


```
RESULT 2
ID N81164 standard; DNA: 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KM E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key
FT misc_feature
FT 19..69
FT /tag= a
FT /function-multiple cloning site
FT primer_bind
FT 187..204
FT /tag= b
EP-285123-A.
PN
PD 05-MAY-1988.
PE 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI: 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P: English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match
Best Local Similarity 6.1%; Score 43; DB 1; Length 204;
Matches 7; Conservative 56; Mismatches 34; Indels 0; Gaps 0;

DB 92 thyrtrmbvryrdsdaaawycyrrsvkydcynachdhdyvbbvynvnhmc 151
QY 320 CCCGGCACCCACGCTGCTGGACCTGGTGAATGCACCTGTGCGGAGAGTG 379
DB 152 ncccbnhvchvbnhrmwayrhdardvbcv 188
QY 380 GAGGCTACGCCCTTCAGCCGCGAGACGAGACGCTCC 416

RESULT 3
ID N81164 standard; DNA: 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KM E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key
FT misc_feature
FT 19..69
FT /tag= a
FT /function-multiple cloning site
FT primer_bind
FT 187..204
FT /tag= b
EP-285123-A.
PN
PD 05-MAY-1988.
PE 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI: 88-279927/40.
```

```
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P: English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match
Best Local Similarity 6.1%; Score 43; DB 1; Length 204;
Matches 7; Conservative 52; Mismatches 30; Indels 0; Gaps 0;

DB 90 thyrtrmbvryrdsdaaawycyrrsvkydcynachdhdyvbbvynvnhn 149
CP 350 TTCACACAGTTCACAGACGCTGCGGCGCTGTGCTGCACAGCCACAGCCCTCA 291
DB 150 ncccbnhvchvbnhrmwayrhd 178
CP 290 CTCGCGCTCTCCAGAGAGATGATGTT 262

RESULT 4
ID 051746 standard; cDNA: 91 BP.
AC 051746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KM Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PE 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3, Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (051735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also 051735-45 and 051747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match
Best Local Similarity 5.8%; Score 41; DB 9; Length 91;
Matches 0; Conservative 44; Mismatches 3; Indels 0; Gaps 0;

DB 14 hsyvrvhvhshhsvhvhvsvvvhvvhvvhvvhvvhvsv 60
QY 179 TCCCGAGGCAAGAACGATCTGTGCGACGATGAATATGCGCTCC 225

RESULT 5
ID 070465 standard; DNA: 114 BP.
AC 070465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KM TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker.
```

| | | | | | | | |
|---------|--|--|------------|------------|------------|----------|-----|
| KW | direct: | rapid; | detection; | screening; | treatment; | generic; | ss. |
| KS | Synthetic. | | | | | | |
| FH | Key | Location/Qualifiers | | | | | |
| FT | misc_feature | 55..60 | | | | | |
| FT | | /tag= a | | | | | |
| FT | | /note= "this sequence represents 'Z'; Z can be a | | | | | |
| FT | | sequence of 6, 9 or 12 nucleotides (see | | | | | |
| FT | | comments)" | | | | | |
| PA | (UNCL) UNIV NORTH CAROLINA. | | | | | | |
| PR | 31-JAN-1994; US-189331. | | | | | | |
| PR | 30-DEC-1993; US-176500. | | | | | | |
| PR | 01-FEB-1993; US-013416. | | | | | | |
| PR | 01-FEB-1994; US-009977. | | | | | | |
| PD | 18-AUG-1994. | | | | | | |
| PT | Identifying proteins or peptide(s) which bind a ligand - by | | | | | | |
| PT | screening a recombinant vector library expressing fusion proteins | | | | | | |
| PS | comprising a binding domain and an effector domain | | | | | | |
| PS | Disclosure; page 35; 255pb; English. | | | | | | |
| CC | 070465 is a generic DNA sequence used to generate random TSAR (Totally | | | | | | |
| CC | Synthetic Affinity Reagents) peptides.This generic formula can also be | | | | | | |
| CC | represented as follows: X(NNB)6(TGC)(NNB)11(CNNB)3Y. X | | | | | | |
| CC | and Y are flanking restriction sites (X is not the same as Y) that are | | | | | | |
| CC | not specified further. Other generic sequences are shown in 070466-68. | | | | | | |
| CC | Other specific peptides generated by these generic sequences are shown i | | | | | | |
| CC | R65151-54. TSARS are concatenated heterofunctional proteins or peptides, | | | | | | |
| CC | comprising at least two functional regions - a binding domain with | | | | | | |
| CC | affinity for a ligand and a second effector peptide portion that is | | | | | | |
| CC | chemically or biologically active.They may further comprise a linker | | | | | | |
| CC | peptide between the 2 domains.The oligonucleotides are also designed so | | | | | | |
| CC | that the expressed peptide contains 2 or 4 cysteine residues positioned | | | | | | |
| CC | in, or flanking, the unpredicted or variant residues. These residues | | | | | | |
| CC | confer some degree of conformational rigidity to the peptides. The TSARS | | | | | | |
| CC | or compms. comprising a TSAR binding domain can be used in vivo to | | | | | | |
| CC | deliver a chemically or biologically active moiety, eg. metal ion, t | | | | | | |
| CC | radioisotope, peptide, toxin or enzyme, to the specific target or on the | | | | | | |
| CC | cell. They can also replace the function of macromolecules, eg. | | | | | | |
| CC | monoclonal or polyclonal antibodies and therefore circumvent the need | | | | | | |
| CC | for complex methods of hybridoma formation or in vivo antibody | | | | | | |
| CC | production. The TSARS are easily formatised and have designed | | | | | | |
| CC | activity allowing direct and rapid detection in a screening process. | | | | | | |
| SO | Sequence 114 BP; 0 A; 2 C; 2 G; 2 T; | | | | | | |
| Dc | Query Match 5.1%; Score 36; DB 12; Length 114; | | | | | | |
| Oy | Best Local Similarity 3.6%; Pred. No. 1.91e-08; | | | | | | |
| Matches | 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0; | | | | | | |
| Dc | 3 bnbnnbnbnbnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnnnnnn 62 | | | | | | |
| Oy | 224 GTCTGTCGTGATPAGCGGGACGCGAAGACTACATCCTCTCGAGAGGAGTGA 293 | | | | | | |
| Dc | 63 bnbnnbn 114 | | | | | | |
| Oy | 294 GGGGCTGCTGGCCTGTGCAACAACGCCGCCGACCCAGCTGCTGTAACCTG 345 | | | | | | |
| RESULT | 6 | | | | | | |
| ID | 070469 standard; DNA: 114 BP. | | | | | | |
| AC | 070469; | | | | | | |
| DT | 07-APR-1995 (first entry) | | | | | | |
| DE | Generic DNA sequence to generate a random TSAR peptide library. | | | | | | |
| KW | TSAR: totally synthetic affinity reagent; synthetic; binding domain; | | | | | | |
| KW | effector domain; concatenated heterofunctional protein; linker; | | | | | | |
| OS | direct; rapid; detection; screening; treatment; generic; ss. | | | | | | |
| FH | Key | Location/Qualifiers | | | | | |
| FT | misc_feature | 55..60 | | | | | |
| FT | | /tag= a | | | | | |
| FT | | /note= "this sequence represents 'Z'; Z can be a | | | | | |
| FT | | sequence of 6, 9 or 12 nucleotides (see | | | | | |
| FT | | comments)" | | | | | |

```

PT  MO9418318-A.
PD  18-AUG-1994.
PF  01-FEB-1994;  U00977.
PR  01-FEB-1993;  US-013416.
PR  30-DEC-1993;  US-176500.
PR  31-JAN-1994;  US-189331.
PA  (UNYC-) UNIV NORTH CAROLINA.
PI  Fowlkes DM, Kay BK.
PI  WPI: 94-279739/34.
PT  Identifying proteins or peptide(s) which bind a ligand - by
PT  screening a recombinant vector library expressing fusion proteins
PT  comprising a binding domain and an effector domain
PS  Disclosure: Page 35; 255pp; English.
CC  Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC  CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC  (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC  sites (X is not the same as Y) that are not specified further. This
CC  sequence generates peptides that are cloverleaf in structure. Other
CC  generic sequences are shown in Q70465-68. Other specific peptides
CC  generated by these generic sequences are shown in R65150-54. TSARs are
CC  concatenated heterofunctional proteins or peptides, comprising at least
CC  two functional regions - a binding domain with affinity for a ligand and
CC  a second effector peptide portion that is chemically or biologically
CC  active. They may further comprise a linker peptide between the 2 domains.
CC  The oligonucleotides are also designed so that the expressed peptide
CC  contains 2 or 4 cysteine residues positioned in, or flanking, the
CC  conformational rigidity to the peptides. The TSARs or compsns. comprising
CC  informational rigidity to the peptides. The TSARs or compsns. comprising
CC  a TSAR binding domain can be used in vivo to deliver a chemically or
CC  biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC  or enzyme, to the specific target or on the cell. They can also replace
CC  the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC  and therefore circumvent the need for complex methods of hybridoma
CC  formation or in vivo antibody production. The TSARs are easily
CC  characterised and have designed activity allowing direct and rapid
CC  detection in a screening process.
SQ  Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;
SQ  Query Match 5.1%; Score 36; DB 12; Length 114;
SQ  Best Local Similarity 6.6%; Pred. No. 1.91e-08;
SQ  Matches 7; Conservative 31; Mismatches 68; Indels 0; Gaps 0;

Db 9 bnnbnnbnnbnnbnnbnnbntgcnnbnnbnnbnnbnnbnnnnnnbnnbntg 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 231 GGGGCTCTGTGGATAGCGGAGCTCGAACAATCACTACTCTCGAGAGCGGAG 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 cnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 291 TGAGGGCTGTGTGGCTGTGGACCAACGCCGCGACCCGAGCTGC 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
DT 070467 standard; DNA: 114 BP.
AC Q70467;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FT Key Location/Qualifiers
FT misc.feature 55..60
FT FT /*tag= a
FT FT /note= "this sequence represents 'Z', Z can be a
FT FT sequence of 6, 9 or 12 nucleotides (see
FT FT comments)"
FT FT
FT PN MO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR 30-DEC-1993; US-013416.
PR 31-JAN-1994; US-189331.

```

[illegible]


```

FT      sequence of 6, 9 or 12 nucleotides (see
FT      comments)"
PN      W09418318-A.
PD      18-AUG-1994.
PF      01-FEB-1994.
PR      01-FEB-1993; US-013416.
PR      30-DEC-1993; US-176500.
PR      31-JAN-1994; US-189331.
PA      (UYNC-) UNIV NORTH CAROLINA.
PI      Fowlkes DM, Kay BK.
DR      WPI: 94-279739/34.
DR      P-PSDB: R65150 and R65151.
PT      Identifying proteins or peptide(s) which bind a ligand - by
PT      screening a recombinant vector library expressing fusion proteins
PT      comprising a binding domain and an effector domain
PS      Disclosure; Page 35; 255pp; Eng11sh.
CC      Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC      Synthetic Affinity Reagents) peptides. This generic formula can also be
CC      represented as follows: X(NNB)6(TGC) (NNB)11Z(NNB)14(TGC) (NNB)3Y. X
CC      and Y are flanking restriction sites (X is not the same as Y) that are
CC      not specified further. Other generic sequences are shown in Q70465-68.
CC      Other specific peptides generated by these generic sequences are shown in
CC      R65151-51. TSARs are concatenated heterofunctional proteins or peptides,
CC      comprising at least two functional regions - a binding domain with
CC      affinity for a ligand and a second effector peptide portion that is
CC      chemically or biologically active. They may further comprise a linker
CC      peptide between the 2 domains. The oligonucleotides are also designed so
CC      that the expressed peptide contains 2 or 4 cysteine residues positioned
CC      in, or flanking, the unpredicted or variant residues. These residues
CC      confer some degree of conformational rigidity to the peptides. The TSARs
CC      or compms, comprising a TSAR binding domain can be used in vivo to
CC      deliver a chemically or biologically active moiety, eg. metal ion,
CC      radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC      cell. They can also replace the function of macromolecules, eg.
CC      monoclonal or polyclonal antibodies and therefore circumvent the need
CC      for complex methods of hybridoma formation or in vivo antibody
CC      production. The TSARs are easily characterised and have designed
CC      activity allowing direct and rapid detection in a screening process.
SQ      Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 4.7%; Score 33; DB 12; Length 114;
Best Local Similarity 4.6%; Pred. No. 1,30e-06;
Matches 5; Conservative 31; Mismatches 73; Indels 0; Gaps 0;

Dn 6 bnnbnnbnnbntgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 65
Cp 423 GAACACAGGAGCGTTCGTCGGCTGAAGGGGGGAGTAGCTCCACATCGCCAAAGTG 364
Dn 66 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnn 114
Cp 363 GCACACAGTGCAATGCACGAGTTCACGACGACGCGGGGTGCCGGCGGTG 315

RESULT 13
ID 070471 standard; DNA: 114 BP.
AC 070471;
DT 10-APR-1995 (first entry)
KW Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; genetic; ss.
OS Synthetic.
FS key Location/Qualifiers
FH misc_feature 55..60
FT /*tag= a
FT /note="encoded by Z (see comments)"
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; US00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.

```

```
PI Fowlkes' M., Kay BK;
DR WPt.: 94-279739/34.
P-PDSB: R58379.
PT Identifying proteins or peptide(s) which bind a ligand - by
PS screening a recombinant vector library expressing fusion proteins
PR comprising a binding domain and an effector domain
PS Disclosure: Page 36; 255pp: English.
CC Q70477 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(CAC)(NNB)9(CAC)(NNB)2(CAC)(NNB)4(CAC)2
CC -(NNB)6(CAC)(NNN)(CAC)(NNB)2v. x and y are flanking restriction sites
CC (x is not the same as Y) that are not specified further. The peptides
CC generated by this and other generic sequences (Q70470-73) have invariant
CC histidine residues incorporated into variant sequences. TSARS are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and/or
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains
CC The TSARS or compns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety. eg. metal
CC ion, radiolotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARS are easily characterised and have designed
CC actively allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 7 A; 14 C; 0 G; 0 T;
```

```
Db Query Match 4.2% Score 30; DB 12; Length 114;
Best Local Similarity 9.8%; Pred. No. 7.83e-05;
Matches 11; Conservative 29; Mismatches 72; Indels 0; Gaps 0.
```

```
Dc 3 bnnbnbnbnbnbdacnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnnaaa 62
Oy 504 CGGGGCGAGGTAGTGCTGCATCACCAGTGAATACTCATGCAGAACCACAGTTCAATCAA 563
      :   :   :   ||| |   :   :   :   :   :   :   :   :   :   !
Oy 564 AGCCACATCTGTCGCCAAGAACGAGGCTTAGAGACAATATCTACTCTTC 615
```

```
Db 63 cnbdbnbnbnbdbaccanbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbb 114
      :   :   :   ||| |   :   :   :   :   :   :   :   :   :
```

```
OY RESULT 14
ID O70472 standard; DNA; 114 BP.
AC O70472:
DT 10-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; genetic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60 /*tag= a /note= "encoded by Z"
```

```
PN W09418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994: U00977.
PR 01-FEB-1993: US-013416.
PR 30-DEC-1993: US-176500.
PR 31-JAN-1994: US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPt.: 94-279739/34.
D P-PDSB: R58383.
```

```
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PR comprising a binding domain and an effector domain
PS Disclosure: Page 36; 255pp: English.
CC Q70477 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides .This generic formula can also be
CC represented as follows: X(NNB)1(CAC)(NNB)11(CAC)(NNB)(CAC)(NNB)2Z((NNB)6
CC -(CAC)(NNB)5(CAC)2(NNB)4. x and y are flanking restriction sites
CC (X is not the same as Y) that are not specified further. The peptides
```


This Page Blank (uspto)


```

Db      188  IPIAOMCLNDEGGPSSLSHRSTFLKXVLEEC-DID-GRSY-R-Q-IIHSRTIKTD-ND 241
Oy      259  VSRVAOLCRDQGGESSLSYSKWNFTFLKAWLVSDATATNNFRRLDVLPLDPSCQWRD 318
Db      242  TLIVFEDPSYSKSACTYSMTNTIKOSFYSKLEGYTKOLPSPAPGICLPAGKVSHTTF 301
Oy      319  TRVYGVFSNPWNMSAVCVYSLGDIDNVFRRSSJLKGTHSSLPLNRPCKCLPDQPIETFT 378
Db      302  EYEKYNVLDDITKPLS 318
Oy      379  QVADRHPVAVQVRPEPMG 395

RESULT 2
ID      VA39_VACCV      STANDARD:      PRT:      441 AA.
AC      P24764;
DT      01-MAR-1992 (REL. 21, CREATED)
DI      01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT      01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE      PROTEIN A39.
GN      A39R OR (SALL9R AND SALLFR).
OS      VACCINIA VIRUS (STRAIN WR).
OC      VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC      ORTHOPOXVIRUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91310644.
RA      AMEGADZIE B.Y., AHN B.-Y., MOSS B.;
RT      "Identification, sequence, and expression of the gene encoding a Mr
RT      35,000 subunit of the vaccinia virus DNA-dependent RNA polymerase.",
RL      J. BIOL. CHEM. 266:13712-13718(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91259063.
RA      SMITH G.L., CHAN Y.S., HOWARD S.T.;
RT      "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT      the right inverted terminal repeat.";
RL      J. GEN. VIROL. 72:1349-1376(1991).
CC      -I- CAUTION: IN REF.2 THIS SEQUENCE IS IN TWO PART (SALL9R AND
CC      SALLFR) DUE TO THE PRESENCE OF FRAMESHIFT COMPARED TO THAT OF
CC      REF.1.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as their content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL, M61187; G335798;
DR      EMBL, D11079; G222726; ALT_SEQ.
DR      EMBL, X57318; G63254;
DR      PIR, S29921, S29921.
SO      SEQUENCE 441 AA; 50185 MW; 26823468 CRC32;

Query Match          9.3%; Score 463; DB 1; Length 441;
Best Local Similarity 33.9%; Pred. No.2, OF=81;
Matches 87; Conservative 57; Mismatches 96; Indels 17; Gaps 16;

Db      115  LVCGTNGNPFCKWK-IDGSDDPKRRHGRGAPYQNSKVTIISHNGC-VLSDINISK-EG-I 170
Oy      141  LACGTNANHPSQCMNLNVGVVPLGLGMRGAPSPDENSLVLEFGDEYVSTRKQYENKI 200
Db      171  KRMREFDPCGYDLYTADNVIPKDG-LGAPVFDKGTID-KYIILFTTIGSKR--I-VK 225
Oy      201  PRFRIRRE-S-ELTSTVWQNPQFIKATIVHOQAAYDDKRYFFREDNDPKNEAPLN 258
Db      226  IPIYIOMCLNDEGGPSSLSHRSTFLKXVLEEC-DID-GRSY-R-Q-IIHSRTIKTD-ND 279
Oy      259  VSRVAOLCRDQGGESSLSYSKWNFTFLKAWLVSDATATNNFRRLDVLPLDPSCQWRD 318

```

```

Db      280  TTVTFEDSDSALCTSMNTIKSEFSKTEGTYTKOLPSPASGICLPACNVPHHTF 339
QY      319  TRYGVGSFNPMWNSACVYSLGDIDKVFRTISLKLGHSSLPNRPBGKCLPDQPIPTETF 378
Db      340  EVIEKYNVLDIDIKPLS 356
QY      379  QVADRHPVEVAQAEVPMG 395

RESULT      3
ID          SEX_HUMAN      STANDARD:      PRT;    1871 AA.
AC          PS1805;
DT          01-OCT-1996 (REL. 34, CREATED)
DD          01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT          01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE          TRANSMEMBRANE PROTEIN SEX PRECURSOR.
GN          SEX.
OS          HOMO SAPIENS (HUMAN).
OC          EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC          PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN          [1]
RP          SEQUENCE FROM N.A.
RC          TISSUE=FETAL BRAIN, SKELETAL MUSCLE, AND EMBRYO;
RX          MEDLINE; 96149362.
RA          MAESTRINI E., TAMAGNONE L., LONGATI P., CREMONA O., GULISANO M.,
RA          BIONE S., TAMANINI F., NEEL B.G., TOMILO D., CONGOLIO P.M.;
RT          "A family of transmembrane proteins with homology to the
RT          MET/hepatocyte growth factor receptor.";
RL          PROC. NATL. ACAD. SCI. U.S.A. 93:674-678(1996).
CC          -!- FUNCTION: PUTATIVE RECEPTOR INVOLVED IN THE DEVELOPMENT OF NEURAL
CC          AND EPITHELIAL TISSUES.
CC          -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC          -!- TISSUE SPECIFICITY: PREDOMINANT IN NEURAL TISSUE DURING
CC          DEVELOPMENT (BY SIMILARITY); WIDELY EXPRESSED IN ADULT TISSUE.
CC          -!- SIMILARITY: WITH THE EXTRACELLULAR DOMAIN OF THE MET/ROM/HGF
CC          RECEPTORS AND THE CYTOPLASMIC DOMAIN OF SEP, NOV AND OCT.
-----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
-----
DR          EMBL; X87852; EI83847; -.
KW          MIN; 300022; -.
KR          RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT          SIGNAL              1           19      POTENTIAL.
FT          CHAIN               20        1871     TRANSMEMBRANE PROTEIN SEX.
FT          DOMAIN             20       1220     EXTRACELLULAR (POTENTIAL).
FT          TRANSMEM          1221      1241     POTENTIAL.
FT          DOMAIN            1242      1871     CYTOPLASMIC (POTENTIAL).
FT          CARBOHYD           59         59      POTENTIAL.
FT          CARBOHYD          548        548      POTENTIAL.
FT          CARBOHYD          637        637      POTENTIAL.
FT          CARBOHYD          738        738      POTENTIAL.
FT          CARBOHYD          746        746      POTENTIAL.
FT          CARBOHYD          1009       1009      POTENTIAL.
FT          CARBOHYD          1036       1036      POTENTIAL.
FT          CARBOHYD          1073       1073      POTENTIAL.
FT          CARBOHYD          1115       1115      POTENTIAL.
FT          CARBOHYD          1162       1162      POTENTIAL.
SO          SEQUENCE          1871 AA;   207661 MW; 9A11046A CRC32;

Query Match      3.3%; Score 163; DB 1; Length 1871;
Best Local Similarity 29.5%; Pred.No. 5,75e-12;
Matches      26; Conservative      29; Mismatches 29; Indels 4; Gaps 3;

```

| ID | 4 | STANDARD: | PT: | 533 AA. |
|--------|--|-------------------------------|-----|---------|
| DB | 484 | SOLPVENTCEYOYS--CAACLGSDGDCGM | 510 | |
| OY | 486 | SOVPLDCEYGGGCGHGLMSRDPYCGM | 513 | |
| RESULT | | | | |
| ID | ARSB_HUMAN | STANDARD: | PT: | 533 AA. |
| AC | P15848: | | | |
| DT | 01-APR-1990 (REL. 14, CREATED) | | | |
| DT | 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE) | | | |
| DT | 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE) | | | |
| DE | ARLSULFATASE B PRECURSOR (EC 3.1.6.12) (ASB) (N-ACETYLGLACTOSAMINE-4-SULFATASE) (G4S). | | | |
| CN | ARSB. | | | |
| OS | HOMO SAPIENS (HUMAN). | | | |
| OC | EDUARVOT: METAQO: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: | | | |
| OC | PRIMATES: CATARRHINI: HOMINIDAE: HOMO. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. | | | |
| EX | MEDLINE: 90153994. | | | |
| RA | PETERS C., SCHMIDT B., ROMERSKIRCH W., RUPP K., ZUEHLSDORF M., | | | |
| RA | VINGRON M., MEYER H.E., POHLMANN R., VON FIGURA K.; | | | |
| RT | "Phylogenetic conservation of arylsulfatases. cDNA cloning and | | | |
| RT | expression of human arylsulfatase B."; | | | |
| RL | J. BIOL. CHEM. 265:3374-3381(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | MEDLINE: 90152677. | | | |
| RA | SCHUCHMAN E.H., JACKSON C.E., DESNICK R.J.; | | | |
| RT | "Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full- | | | |
| RT | length cDNA, and regions of amino acid identity with arylsulfatases A | | | |
| RL | and C."; | | | |
| RL | GENOMICS 6:149-158(1990). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | MEDLINE: 93332648. | | | |
| RA | MODARESSI S., RUPP K., VON FIGURA K., PETERS C.; | | | |
| RT | "Structure of the human arylsulfatase B gene."; | | | |
| RL | BIOL. CHEM. HOPPE-SEYLER 374:327-335(1993). | | | |
| RN | [4] | | | |
| RP | SEQUENCE OF 1-104 FROM N.A. | | | |
| EX | MEDLINE: 92028992. | | | |
| RA | LITVENS T., MORRIS C.P., GIBSON G.J., BECKMANN K.R., HOPWOOD J.J.; | | | |
| RT | "Human N-acetylglactosamine-4-sulphatase: protein maturation and | | | |
| RL | isolation of genomic clones."; | | | |
| RL | BIOCHEM. INT. 24:209-215(1991). | | | |
| RN | [5] | | | |
| RP | 2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE. | | | |
| EX | MEDLINE: 95354208. | | | |
| RA | SCHMIDT B., SELMER T., INGENDOH A., VON FIGURA K.; | | | |
| RT | "A novel amino acid modification in sulfatases that is defective in | | | |
| RT | multiple sulfatase deficiency."; | | | |
| RL | CELL 82:271-278(1995). | | | |
| RN | [6] | | | |
| RP | X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS). | | | |
| EX | MEDLINE: 97184692. | | | |
| RA | BOND C.S., CLEMENTS P.R., ASHBY S.J., COLLYER C.A., HARROP S.J., | | | |
| RA | HOPWOOD J.J., GUSS J.M.; | | | |
| RT | "Structure of a human lysosomal sulfatase."; | | | |
| RL | STRUCTURE 5:277-289(1997). | | | |
| RN | [7] | | | |
| RP | VARIANTS MPS-VI ARG-117: PRO-236 AND TYR-405. | | | |
| EX | MEDLINE: 92197625. | | | |
| RA | JIN W.-D., JACKSON C.E., DESNICK R.J., SCHUCHMAN E.H.; | | | |
| RT | "Mucopolysaccharidosis type VI: identification of three mutations in | | | |
| RT | the arylsulfatase B gene of patients with the severe and mild | | | |
| RL | phenotypes provides molecular evidence for genetic heterogeneity."; | | | |
| RL | AM. J. HUM. GENET. 50:795-800(1992). | | | |
| RN | [8] | | | |
| RP | VARIANT MPS-VI GLY-137, AND VARIANT MET-376. | | | |
| EX | MEDLINE: 92042029. | | | |
| RA | WICKER G., PRILL V., BROOKS D., GIBSON G., HOPWOOD J., | | | |
| RA | VON FIGURA K., PETERS C.; | | | |

RA "Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An intermediate
 RT clinical phenotype caused by substitution of valine for glycine at
 RT position 137 of arylsulfatase B.";
 RL J. BIOL. CHEM. 266:21386-21391(1991).
 RN [9]
 RP VARIANTS MPS-VI MEN-92; GLN-95; CYS-210; PRO-393 AND PRO-498.
 RX MEDLINE: 96213747.
 RA LITJENS T., BROOKS D.A., PETERS C., GIBSON G.J., HOPWOOD J.J.:
 RT "Identification, expression, and biochemical characterization of N-
 RT acetylgalactosamine-4-sulfatase mutations and relationship with
 RT clinical phenotype in MPS-VI patients.";
 RL AM. J. HUM. GENET. 58:1127-1134(1996).
 RN [10]
 RP VARIANTS MPS-VI TRP-152 AND GLN-160.
 RX MEDLINE: 94171224.
 RA VOSKOBOEVA E., ISBRANDT D., VON FIGURA K., KRASNOPOLSKAYA X.,
 RA PETERS C.:
 RT "Pur novel mutant alleles of the arylsulfatase B gene in two
 RT patients with intermediate form of mucopolysaccharidosis VI
 RT (Maroteaux-Lamy syndrome).";
 RL HUM. GENET. 93:259-264(1994).
 RN [11]
 RP VARIANT MPS-IV ARG-302.
 RX VILLANI G.R.D., BALZANO N., DI NATALE P.:
 RA "Two novel mutations of the arylsulfatase B gene in two Italian
 RT patients with severe form of mucopolysaccharidosis.";
 RL HUM. MUTAT. 11:410-410(1998).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 4-SULFATE GROUPS OF THE
 CC N-ACETYL-D-GLYCOSAMINE 4-SULFATE UNITS OF CHONDROITIN SULFATE
 CC AND DERMATAN SULFATE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
 CC SYNDROME, ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
 CC THIS DISEASE IS CHARACTERIZED BY THE ACCUMULATION OF DERMATAN
 CC SULFATE IN LYSOSOMES. CLINICAL FEATURES CAN INCLUDE ABNORMAL
 CC GROWTH, SHORT STATURE, STIFF JOINTS, SKELETAL MALFORMATIONS,
 CC CORNEAL CLOUDING, HEPATOSPLENOMEGALY, AND CARDIAC ABNORMALITIES.
 CC -1- A WIDE VARIATION IN CLINICAL SEVERITY IS OBSERVED.
 CC -1- DISEASE: MULTIPLE SULFATASE DEFICIENCY (MSD) IS A DISORDER THAT
 CC COMBINES FEATURES OF METACHROMATIC LEUKODYSTROPHY AND OF
 CC MUCOPOLYSACCHARIDOSIS. IT IS CHARACTERIZED BY A DECREASED
 CC ACTIVITY OF ALL KNOWN SULFATASES. IT SEEMS TO BE CAUSED FROM THE
 CC LACK OF POST-TRANSLATIONAL MODIFICATION OF A CYSTEINE INTO 2-
 CC AMINO-3-OXOPROPIONIC ACID.
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J05225; G19077; -.
 DR EMBL: M33373; G19030; -.
 DR EMBL: X72735; G825628; -.
 DR EMBL: X72736; G825628; JOINED.
 DR EMBL: X72737; G825628; JOINED.
 DR EMBL: X72738; G825628; JOINED.
 DR EMBL: X72739; G825628; JOINED.
 DR EMBL: X72740; G825628; JOINED.
 DR EMBL: X72741; G825628; JOINED.
 DR EMBL: X72742; G825628; JOINED.
 DR EMBL: S57777; G236698; -.
 DR PIR: A35078; A35078.
 DR PDB: 1FSU; 04-FEB-98.
 DR MIM: 253200; -.
 DR MIM: 272200; -.
 DR PROSITE, PS00523; SULFATASE_1; 1.
 DR PROSITE, PS00149; SULFATASE_2; 1.
 DR PFM: PF00884; Sulfatase; 1.
 DR

| Query Match | Similarity | Score | DB | Length |
|--------------------------|---|--|------|---------|
| Best Local | 41.3% | Pred. No. 4,59e-02 | | |
| Matches 26; Conservative | 13; Mismatches 19; Indels 5; Gaps 4; | | | |
| Db | 4 | RGASLPRG--PGR-RLLPVPLLLLPAPGSGAGSRPHVLL-ADDCGMD | 59 | |
| Qy | 8 | RAAPSAPRAYVGPAPRLGLRLLLLLW-AAASNOGHLRSGPRIFAWKGVQDGR | 66 | |
| Db | 60 | VGF 62 | | |
| Qy | 67 | VDF 69 | | |
| RESULT | 5 | STANDARD: | PRT: | 535 AA. |
| ID | ARSB_FELCA | | | |
| AC | P33727 | | | |
| DT | 01-FEB-1994 (REL. 28, CREATED) | | | |
| DT | 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE) | | | |
| DT | 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE) | | | |
| DE | ARLSULFATASE B PRECURSOR (EC 3.1.6.12) (ASB) (N-ACETYLGALACTOSAMINE-4-SULFATASE) (GAS). | | | |
| GN | ARSB. | | | |
| OS | FELIS SILVESTRIS CATUS (CAT). | | | |
| OC | EDUARDOI; METAQO; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; | | | |
| OC | CARNIVORA; FISSIPEDIA; FELIDAE; FELIS. | | | |
| NC | [1] | | | |
| NC | SEQUENCE FROM N.A. | | | |
| RP | TISSUE-LIVER. | | | |
| RX | MEDLINE; 93052342. | | | |
| RA | JACKSON C.E., YUHKI N., DESNICK R.J., HASKINS M.E., O'BRIEN S.J., | | | |
| RA | SCHUCHMAN E.H.; | | | |
| RT | "Feline arylsulfatase B (ARSB): isolation and expression of the cDNA, | | | |
| RT | comparison with human ARSB, and gene localization to feline | | | |
| RT | chromosome A1." | | | |
| RL | GENOMICS 14:403-411(1992). | | | |
| CC | -1- CATALYTIC ACTIVITY: HYDROLAYSIS OF THE 4-SULFATE GROUPS OF THE | | | |
| CC | N-ACETYL-D-GALACTOSAMINE 4-SULFATE UNITS OF CHONDROITIN SULFATE | | | |
| CC | AND DERMATAN SULFATE. | | | |
| CC | -1- SUBUNIT: HOMODIMER. | | | |
| CC | -1- SUBCELLULAR LOCATION: LYSOSOMAL. | | | |
| CC | -1- DISEASE DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY | | | |

```

CC SYNDROME: ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
CC MPS-VI HAS BEEN DESCRIBED IN SIAMESE CATS.
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/isb-slb.ch).
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: S48472; G258856; -.
DR PIR: A44475; A44475.
DR PROSITE: PS00523; SULFATASE_1; 1.
DR PROSITE: PS00149; SULFATASE_2; 1.
DR PFAM: PF00884; Sulfatase; 1.
DR HSSP: P15848; IFSU.
KW HYDROLASE; SIGNAL; GLYCOPROTEIN; LYSOSOME; MUCOPOLYSACCHARIDOSIS.
FT SIGNAL 1
FT CHAIN 41
FT ACCT_SITE 42 535
FT MOD_RES 93 93
FT DISULFID 119 523
FT DISULFID 123 157
FT DISULFID 183 194
FT DISULFID 407 449
FT CARBOHYD 190 190
FT CARBOHYD 281 281
FT CARBOHYD 293 293
FT CARBOHYD 428 428
FT CARBOHYD 460 460
FT SEQUENCE 535 AA; 59753 MW; 4C473698 CRC32;
SQ
Query Match 2.2%; Score 110; DB 1; Length 535;
Best Local Similarity 34.6%; Pred. No. 1,58e+02;
Matches 26; Conservative 19; Mismatches 31; Indels 3; Gaps 3;

Db 4 RGAASLPKPGPPRRRLLELGVLLRLRLRLPSPGAGAGAD-RPPLHYE-VLADDLGMD 61
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 8 RAAPSAPAPARVPGPARG-LPLRLRLLLMAAASAGHLRSPRIFAVWKGHVQDR 66
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 62 VSFHGSNIRPHDELAAAGV 82
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 67 VDFGQTEPHTVLFHFGSSSV 87
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 6
AC US08_HCMVA STANDARD; PRT; 227 AA.
AD P09730;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN HXLf4 PRECURSOR.
GN US8.
OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).
OC VIRUSES: DSDNA VIRUSES, NO RNA STAIN; HERPESVIRIDAE:
OC BETAHERPESVIRINAE; CYTOMEGALOVIRUS.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87169717.
RA WESTON K., BARRELL B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. MOL. BIOL. 192:177-208(1986).
[2]
RP COMPLETE GENOME.
RX MEDLINE: 90269039.
RA CHEE M.S., BANKER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R.,
RA HORSNELL T., HUTCHISON C.A. III, KOUZARIDES T., MARTIGNETTI J.A.,
RA PRADDE E., SATCHEWELL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;
RT "Analysis of the protein-coding content of the sequence of human

```



```

RT      adenoocarcinoma cells transformed with rat T cell alloantigen Rf6.2." ;
RJ      J. BIOL. CHEM. 269:9420-9423(1994) .
CC      -1- FUNCTION: HAS BOTH NAD+ GLYCOHYDROLASE AND ADP-RIBOSYLTRANSFERASE
CC      ACTIVITY (TO A LESSER EXTENT) .
CC      -1- CATALYTIC ACTIVITY: NAD(+) + L-ARGININE = NICOTINAMIDE +
CC      N2-(ADP-D-RIBOSYL)-L-ARGININE .
CC      -1- CATALYTIC ACTIVITY: NAD(+) + H2O = NICOTINAMIDE + ADP-RIBOSE .
CC      -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR .
CC      -1- TISSUE SPECIFICITY: POSTTHYMIC T CELLS .
CC      -1- PIM: TWO DISULFIDE BONDS ARE PRESENT (PROBABLE) .
CC      -1- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE
CC      FAMILY .
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.sib.ch) .
-----
DR      EMBL; M85193; G206806; -
DR      EMBL; X99123; E257753; -
DR      EMBL; X99122; E257751; -
DR      PIR: A34866; A34866.
DR      PROSITE; PS01291; ART: 1.
DR      PFAM; PF01129; ART: 1.
DR      TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; NAD; SIGNAL;
DR      T-CELL DIFFERENTIATION; GPI-ANCHOR .
FT      SIGNAL          1          20
FT      CHAIN           21          246
FT      PROPEP          247          275
FT                                     T-CELL, ECTO-ADP-RIBOSYLTRANSFERASE 2.
FT      LIPID           246          246      HYDROPHOBIC, REMOVED DURING MATURATION
FT      ACT_SITE        209          209      GPI-ANCHOR (BY SIMILARITY) .
FT      ACT_SITE        209          209      BY SIMILARITY .
SQ      SEQUENCE        275 AA;  31438 MW;  DBEAB84E CRC32;
Query Match      2.1%; Score 102; DB 1; Length 275;
Best Local Similarity 30.6%; Pred. No. 2.57e-01;
Matches 19; Conservative 17; Mismatches 21; Indels 5; Gaps 5;
Db 194 LGVYKKEE-RRPOEE-VLIRGYEVQKVTGQNEFLDSPKKRSNYNCVSSAGAR 251
QY 163 LGE-MRGTAAPSDDNSLVLEEGDEVISTIRKQYNGK-IPFRRIKGESE-LVTSPIYV 219
Db 252 ES 253
QY 220 QN 221

```

```

SQ      SEQUENCE      141 AA: 15745 MW: 6327FEAL CRC32:
Query Match                               2.0%: Score 98; DB 1; Length 141;
Best Local Similarity 33.3%: Pred. No. 9,70e-01;
Matches 17; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

Db      39 PPPPPPPPPPPPPPPPPPPRRRAHHNNHPLFLFFKKSYSNNHMCGR 89
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
OY      3 PPPPGRAAPSAPRARVPGPPARGLPLRLRLLLMLMAAASAGCHLSGPR 53

RESULT 9
ID      TNFB, RABBIT      STANDARD;      PRT;      197 AA.
AC      P10154;
DT      01-MAR-1989 (REL. 10, CREATED)
DT      01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE      LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN      LTA OR TNFB.
OS      ORYCTOLAGUS CUNICULUS (RABBIT).
OC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC      LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 91065534.
RA      SHAHNOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
RA      NEDOSPASOV S.A.:
RT      "Structural analysis of the rabbit TNF locus, containing the genes
RT      encoding TNF-Delta (lymphotoxin) and TNF-alpha (tumor necrosis
RT      factor).";
RL      GENE 95:215-221 (1990).
RN      (2)
RP      SEQUENCE FROM N.A.
RX      MEDLINE: 90220566.
RA      SHAHNOV A.N., KUPRASH D.V., TURETSKAYA R.L., AZIZOV M.M.,
RA      ANDEREYEVA A.V., NEDOSPASOV S.A.:
RT      "Cloning and structural analysis of genes coding for tumor necrosis
RT      factor and lymphotoxin in rabbits.";
RL      MOL. BIOL. (MOSK) 23:1743-1750(1989).
CC      -I- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC      IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC      -I- SUBUNIT: HETERODIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC      OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC      -I- SUBCELLULAR LOCATION: SECRETED.
CC      -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X55745; G297168; ALT_SEQ.
DR      EMBL: M60340; G165753; -.
DR      EMBL: M60341; G165758; -.
DR      PIR: JH0309; JH0309.
DR      PIR: PNO098; PNO098.
DR      PROSITE: PS00251; TNF_1; 1.
DR      PROSITE: PS50049; TNF_2; 1.
DR      PFMAM; PF00229; TNF; 1.
KW      CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
FT      SIGNAL      1      26
FT      CHAIN      27      197      LYMPHOTOXIN-ALPHA.
FT      CARBOHYD      88      88      POTENTIAL.
SO      SEQUENCE      197 AA: 21126 MW: 38427DBB CRC32;

Query Match                               2.0%: Score 99; DB 1; Length 197;
Best Local Similarity 33.3%: Pred. No. 6,99e-01;
Matches 31; Conservative 28; Mismatches 22; Indels 12; Gaps 11;
3 PPGRIYIDL-L-VGLLLAPPPGAOG-L-PGAE-FPPSAARNAOQRLQKHGHSTIKPAA 57

```

| | | |
|--------|---|---|
| Qy | 21 | PARLGLPRLRLILLIMAAASGHLRSPRIFAWKGHVGDQD--FQD-T-EPT 76 |
| Db | 58 | HLVGDPSAODSLMRANTDRAFLRHGFSLSNNS 90 |
| Qy | 77 | VLFHPDS-SSV-WVGGRKVL-FDFPESKNA 106 |
| RESULT | 10 | |
| ID | NRT1_RAT | STANDARD: PRI: 275 AA. |
| AC | P17982; | |
| DT | 01-NOV-1990 (REL. 16, CREATED) | |
| DT | 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE) | |
| DT | 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE) | |
| DE | T-CELL ECTO-ADP-RIBOSYLTRANSFERASE 1 PRECURSOR (EC 2.4.2.31) (T-CELL | |
| DE | NAD(P)(+)--ARGININE ADP-RIBOSYLTRANSFERASE 1 (T-CELL MONO(ADP- | |
| DE | RIBOSYL)TRANSFERASE 1) (ALLOANTIGEN RT6.1) (T-CELL SURFACE PROTEIN | |
| GN | RT6.1) | |
| OS | ART2A OR RT6-A. | |
| OS | RATTUS NORVEGICUS (RAT). | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; | |
| OC | RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; RATTUS. | |
| RP | [1] | |
| RC | SEQUENCE FROM N.A. | |
| RC | STRAIN-LEWIS A. | |
| RX | MEDLINE: 90192088. | |
| RA | HAAG F., KOCH F., THIELE H. -G.; | |
| RT | "Nucleotide and deduced amino acid sequence of the rat T-cell | |
| RT | alloantigen RT6.1."; | |
| RT | NUCLEIC ACIDS RES. 18:1047-1047(1990). | |
| RP | [2] | |
| RP | MUTAGENESIS OF GLN-207. | |
| RX | MEDLINE: 96275529. | |
| RA | MAERHA T., HOSHINO S.-I., KATADA T.; | |
| RT | "Increase in ADP-ribosyltransferase activity of rat T lymphocyte | |
| RT | alloantigen RT6.1 by a single amino acid mutation."; | |
| RT | FEBS LETT. 388:189-191(1996). | |
| CC | -1- FUNCTION: HAS NAD+ GLYCOHYDROLASE ACTIVITY AND EXTREMELY LOW ADP- | |
| CC | RIBOSYLTRANSFERASE ACTIVITY. | |
| CC | -1- CATALYTIC ACTIVITY: NAD(+) + L-ARGININE = NICOTINAMIDE + | |
| CC | N2-(ADP-D-RIBOSYL)-L-ARGININE. | |
| CC | -1- CATALYTIC ACTIVITY: NAD(+) + H(2)O = NICOTINAMIDE + ADP-RIBOSE. | |
| CC | -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR. | |
| CC | -1- TISSUE SPECIFICITY: POSTMYC T CELLS. | |
| CC | -1- PM: TWO DISULFIDE BONDS ARE PRESENT (PROBABLE). | |
| CC | -1- SIMILARITY: BELONGS TO THE ARG-SPECIFIC ADP-RIBOSYLTRANSFERASE | |
| CC | FAMILY. | |
| CC | ----- | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | |
| CC | use by non-profit institutions as long as its content is in no way | |
| CC | modified and this statement is not removed. Usage by and for commercial | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch). | |
| CC | or send an email to license@isb-sib.ch). | |
| DR | EMBL: X52082; G57168; -. | |
| DR | EMBL: M31138; G206804; -. | |
| DR | PIR: S08464; S08464. | |
| DR | PROSITE: PS01291; ART; 1. | |
| DR | PEPM: PF01129; ART; 1. | |
| KW | TRANSFERASE: GLYCOSYLTRANSFERASE; GLYCOPROTEIN; NAD; SIGNAL; | |
| FM | T-CELL DIFFERENTIATION; GPI-ANCHOR. | |
| FT | SIGNAL | 1 |
| FT | CHAIN | 20 |
| FT | PROPEP | 21 246 |
| FT | | 247 275 |
| FT | LIPID | 246 246 |
| FT | ACT_SITE | 209 209 |
| FT | CARBOHYD | 58 58 |
| FT | MUTAGEN | 207 207 |
| QO | SEQUENCE | 275 AA; 31388 MW; 52381A84 CRC32; ACTIVITY. |

Query Match 2.0%: Score 99; DB 1; Length 275;
 Best Local Similarity 30.6%; Pred. No. 6.99e-01;
 Matches 19; Conservative 17; Mismatches 21; Indels 5; Gaps 5;

Db 194 LGVIRKES-FYDPOEE-VLIPGYEVOKVTOGNIETLDSPPKRSNNCLYSSAGTR 251
 163 LGE-MRGYAPFSPDENSIVLEFGDEGVSTIRKOEYNGK-IPFRFRINGESE-LYTSPTVM 219

Db 252 ES 253
 220 QN 221

RESULT 11
 ID LMP1_EBV STANDARD; PRT; 386 AA.
 AC P03230;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63) (P25).
 GN BNL1.
 OS EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
 OC VIRUSES; DSDNA VIRUSES; NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 84270667.
 RA BAKER R., BANKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J.,
 RA GIBSON T.J., HARTFILL G., HUDSON G.S., SATCHWELL S.C., SEGWIN C.,
 RA TUFFENL P.S., BARRELL B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL NATURE 310:207-211(1984).
 RN [2]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE: 87112999.
 RA BAICHVAL V.R., SUDGEN B.;
 RT "Posttranslational processing of an Epstein-Barr virus-encoded
 membrane protein expressed in cells transformed by Epstein-Barr
 virus.";
 RL J. VIROL. 61:866-875(1987).
 RN [3]
 RP TRANSFORMING PROPERTIES.
 RX MEDLINE: 88233656.
 RA BAICHVAL V.R., SUDGEN B.;
 RT "Transformation of Balb 3T3 cells by the BNL1-1 gene of Epstein-Barr
 virus.";
 RL ONCOGENE 2:461-467(1988).
 RN [4]
 RP PROCESSING.
 RX MEDLINE: 90112654.
 RA MOORTHY R., THORLEY-LANSON D.A.;
 RT "Processing of the Epstein-Barr virus-encoded latent membrane protein
 p63/LMP.";
 RL J. VIROL. 64:829-837(1990).
 CC -1- FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: V01555; -; NOT_ANNOTATED_CDS.
 DR PIR: A03794; OQB50.
 KM TRANSMEMBRANE: PHOSPHORYLATION; TRANSFORMING PROTEIN.
 FT DOMAIN 1 24 CYTOPLASMIC.
 FT TRANSMEM 25 44 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.

FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT DOMAIN 187 386 CYTOPLASMIC.
 FT PEPTIDE 242 386 P25.
 SQ SEQUENCE 386 AA; 41982 MM; 1F91D6E3 CMC32;

Query Match 2.0%: Score 99; DB 1; Length 386;
 Best Local Similarity 45.5%; Pred. No. 6.99e-01;
 Matches 15; Conservative 9; Mismatches 7; Indels 2; Gaps 2;

Db 7 RGPRGRRPRGPPPLSSGLALLLALLLFW 39
 8 RAAPSAPRA-RVPGPARLGLPLRL-LLW 38

RESULT 12
 ID LMP1_EBVC STANDARD; PRT; 404 AA.
 AC P29362;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63).
 GN BNL1.
 OS EPSTEIN-BARR VIRUS (STRAIN CAO) (HUMAN HERPESVIRUS 4).
 OC VIRUSES; DSDNA VIRUSES; NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92013956.
 RA HU L.F., ZABAROVSKY E.R., CHEN F., CAO S.L., ERNBERG I., KLEIN G.,
 RA WINBERG G.;
 RT "Isolation and sequencing of the Epstein-Barr virus BNL1-1 gene
 (LMP1) from a Chinese nasopharyngeal carcinoma.";
 RL J. GEN. VIROL. 72:2399-2409(1991).
 CC -1- FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X58140; G22938; -
 DR PIR: J01434; LABECA.
 DR PIR: S21660; S21660.
 KM TRANSMEMBRANE: PHOSPHORYLATION; TRANSFORMING PROTEIN.
 FT DOMAIN 1 24 CYTOPLASMIC.
 FT TRANSMEM 25 44 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT DOMAIN 187 404 CYTOPLASMIC.
 SQ SEQUENCE 404 AA; 43769 MM; 154E84C3 CMC32;

Query Match 2.0%: Score 98; DB 1; Length 404;
 Best Local Similarity 50.0%; Pred. No. 9.70e-01;
 Matches 16; Conservative 8; Mismatches 6; Indels 2; Gaps 2;

Db 8 AAPAPRPPLGPPPLSSIGLALLLALLLFW 39
 9 AAPAPRAPVGP-PARLGLPLRL-LLW 38

RESULT 13
 ID SPC3_RAT STANDARD; PRT; 442 AA.
 AC P33671; P97614;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

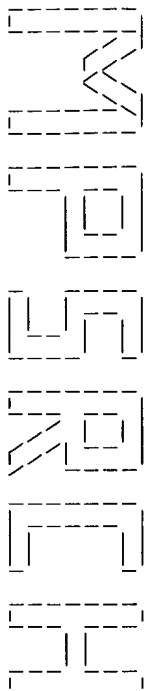
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE 65 KD VIRULENCE PROTEIN.
GN MKAA.
OS SALMONELLA TYPHIMURIUM.
OG PLASMID 96 KB VIRULENCE PEX102.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC SALMONELLA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TWL R66;
RX MEDLINE; 90136009.
RA TAIRA S., RHEN M.;
RT "Identification and genetic analysis of mka -- a gene of the
RT Salmonella typhimurium virulence plasmid necessary for intracellular
RT growth."
RL MICROB. PATHOG. 7:165-173(1989).
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE; 91244158.
RA TAIRA S., BAUMANN M., RIKONEN P., SUKUPOLVI S., RHEN M.;
RT "Amino-terminal sequence analysis of four plasmid-encoded virulence-
RT associated proteins of Salmonella typhimurium."
RL FEMS MICROBIOL. LETT. 61:319-323(1991).
CC -1- FUNCTION: NOT KNOWN. THIS PROTEIN IS INVOLVED IN THE VIRULENCE
CC OF SALMONELLAS.
CC -1- SIMILARITY: TO SIMILAR PROTEIN IN OTHER SALMONELLA VIRULENCE
CC PLASMIDS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z15042; G47783; .
DR PTR: A54540; A54540.
KW PLASMID; VIRULENCE.
FT DOMAIN 367 373 POLY-PRO.
SQ SEQUENCE 591 AA; 65440 MW; B69F25C3 CRC32;

Query Match 2.0%; Score 99; DB 1; Length 591;
Best Local Similarity 23.3%; Pred. No. 6,99e-01;
Matches 21; Conservative 25; Mismatches 41; Indels 3; Gaps 3;

Db 162 HDSNGLHLGLKTAARLSDPOASHTAQW-LVESVTPAGE-HIYSYLAENGDNVDLN 219
QY 124 RDCEVYIILLERSSEGLACGTNARHPCWNVNCTVPLGEMRGYAFSPDENSLVLE 183
Db 220 GNEAGRDRSAMRYLSKV-QYGNATPADLY 248
QY 184 GDEVYSTIRKQYNGKIPFRIRGESELY 213

Search completed: Tue Aug 3 16:29:59 1999
Job time : 61 secs.

This Page Blank (uspto)



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:26:51 1999; Maspar time 28.08 Seconds
950.249 Million cell updates/sec

Tabular output not generated.

Title: >US-09-240-410-2
Description: (1-666) from US09240410.pep
Perfect Score: 4968
Sequence: 1 MTPTTGGRAAPSAPRARVPG.....LAASIMLGVLPTLTGLLVH 666

Scoring table: PAM 150
Gap 11

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r60
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 50.278; Variance 90.656; scale 0.555

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description | Pred. No. |
|------------|-------|---------------|--------|-------|--------------------------------|-----------|
| 1 | 1907 | 38.4 | 653 | 2 | T03102 semaphorin homolog A3 | 0.00e+00 |
| 2 | 737 | 14.8 | 748 | 2 | 148744 semaphorin A - mouse | 8.75e-131 |
| 3 | 725 | 14.6 | 751 | 2 | 148748 semaphorin E - mouse | 4.38e-128 |
| 4 | 716 | 14.4 | 749 | 2 | G01856 semaphorin V - human | 4.61e-126 |
| 5 | 688 | 13.8 | 753 | 2 | G02173 semaphorin III family | 8.77e-120 |
| 6 | 685 | 13.8 | 772 | 2 | A49069 collapsin - chicken | 4.12e-119 |
| 7 | 669 | 13.5 | 772 | 2 | 148747 semaphorin D - mouse | 1.56e-115 |
| 8 | 667 | 13.4 | 666 | 2 | 158169 semaphorin III - mus | 4.37e-115 |
| 9 | 661 | 13.3 | 771 | 2 | D49423 semaphorin C - mouse | 3.58e-114 |
| 10 | 618 | 12.4 | 782 | 2 | 148746 M-sema F protein prec | 5.15e-104 |
| 11 | 522 | 10.5 | 834 | 2 | 566498 semaphorin B - mouse | 7.12e-74 |
| 12 | 480 | 9.7 | 760 | 2 | 148745 A39R protein - vacci | 1.04e-71 |
| 13 | 470 | 9.5 | 403 | 2 | E42521 hypothetical protein | 3.37e-70 |
| 14 | 463 | 9.3 | 441 | 2 | S29921 foscliclin IV precursor | 5.73e-67 |
| 15 | 461 | 9.3 | 730 | 2 | JH0798 semaphorin I precursor | 2.13e-54 |
| 16 | 448 | 9.0 | 1074 | 2 | A59423 semaphorin I - fruit | 9.90e-50 |
| 17 | 389 | 7.8 | 711 | 2 | B49423 semaphorin II precurs | 1.45e-34 |
| 18 | 367 | 7.4 | 656 | 2 | C49423 Sal19R protein - vacc | 7.49e-16 |
| 19 | 362 | 7.3 | 724 | 2 | UC1775 plexin 2 precursor - | 3.14e-11 |
| 20 | 293 | 5.9 | 295 | 2 | JC4975 plexin 1 precursor - | 6.96e-11 |
| 21 | 196 | 3.9 | 1884 | 2 | JC4976 | |
| 22 | 170 | 3.4 | 1872 | 2 | JC4980 | |
| 23 | 168 | 3.4 | 1894 | 2 | JC4980 | |

| | | | | | | |
|----|-----|-----|------|---|--------------------------------|----------|
| 24 | 171 | 3.4 | 1905 | 2 | I51553 Plexin - African claw | 2.11e-11 |
| 25 | 157 | 3.2 | 122 | 2 | J01845 14R protein - variola | 5.17e-09 |
| 26 | 157 | 3.2 | 122 | 2 | H36852 A3R protein - variol | 5.17e-09 |
| 27 | 155 | 3.1 | 142 | 2 | T01776 Salfr protein - vacc | 1.12e-08 |
| 28 | 120 | 2.4 | 775 | 2 | E70320 polyribonucleotide nu | 3.53e-03 |
| 29 | 115 | 2.3 | 422 | 2 | S32357 glial growth factor - | 1.85e-02 |
| 30 | 107 | 2.2 | 533 | 1 | KJH0AB N-acetylglactosamine | 2.37e-01 |
| 31 | 110 | 2.2 | 535 | 2 | A44475 N-acetylglactosamine | 9.24e-02 |
| 32 | 105 | 2.1 | 227 | 2 | S09922 hypothetical protein | 4.39e-01 |
| 33 | 105 | 2.1 | 235 | 1 | Q0BEC9 HXLR protein - human | 4.39e-01 |
| 34 | 102 | 2.1 | 275 | 2 | A34866 conserved surface prote | 1.09e+00 |
| 35 | 103 | 2.1 | 406 | 2 | B69064 conserved hypochetica | 8.05e-01 |
| 36 | 99 | 2.0 | 197 | 2 | JH0309 tumor necrosis factor | 2.64e+00 |
| 37 | 99 | 2.0 | 275 | 2 | S08464 T-cell allonitigen RT | 2.64e+00 |
| 38 | 99 | 2.0 | 366 | 1 | Q0B500 latent membrane prote | 2.64e+00 |
| 39 | 99 | 2.0 | 460 | 2 | S35772 translaton elongatio | 2.64e+00 |
| 40 | 99 | 2.0 | 563 | 2 | S78224 virulence-associated | 2.64e+00 |
| 41 | 99 | 2.0 | 591 | 2 | S26565 virulence-associated | 2.64e+00 |
| 42 | 99 | 2.0 | 597 | 2 | A46050 thyroid/steroid recep | 2.64e+00 |
| 43 | 99 | 2.0 | 1091 | 2 | A58532 glial cell membrane g | 1.97e+00 |
| 44 | 100 | 2.0 | 1338 | 2 | S09982 protein-tyrosine kina | 2.64e+00 |
| 45 | 99 | 2.0 | 1375 | 2 | JC5148 hepatocyte growth fac | 2.64e+00 |

ALIGNMENTS

| RESULT | ENTRY | 1 | ALIGNMENTS |
|-----------------------|---|---|--|
| | T03102 | #type complete | |
| | semaphorin homolog A3 - | alcelaphine herpesvirus 1 | |
| | ORGANISM | #formal_name alcelaphine herpesvirus 1 | |
| | DATE | 24-Mar-1999 #sequence_ | revision 24-Mar-1999 #text_ |
| | | 24-Mar-1999 | change |
| ACCESSIONS | T03102 | | |
| REFERENCE | Z1840 | | |
| | Ensser, A.; Pfanz, R.; Fleckenstein, B. | | |
| | J. Virol. (1997) 71:6517-6525 | | |
| | #title | Primary structure of the alcelaphine herpesvirus 1 genome. | |
| | #accession | T03102 | |
| | ##status | Preliminary; translated from GR/EMBL/DBJ | |
| | ##molecule_type | DNA | |
| | ##residues | 1-653 | ##label ENS |
| | ##cross-references | EMBL:AF005370; NID:g2337967; PID:g2337970 | |
| SUMMARY | #length 653 | #molecular_weight 73645 | #checksum 5501 |
| Query Match | 38.4% | Score 1907; | DB 2; Length 653; |
| Best Local Similarity | 46.2% | Pred. No. 0.00e+00; | |
| Matches | 279; | Conservative | 108; Mismatches 203; Indels 14; Gaps 14; |
| Db | 45 | PAMGTCVCSIRILMIT-SAITAKSRFIDKPRILVLTGFCOHRF-FGPOPHVFLFH 102 | |
| QY | 22 | PARG-LPLRLRLILIMAAASNOGHLKSGPRIFAWKCHVGQDRVFCQTERPHVLFH 80 | |
| Db | 103 | SLNSDDYVGNNTIYLFDFRASHSNASTALINTSTHNRHSSTGENFTLLHNQDTGL 162 | |
| QY | 81 | EPGSSVWVGGRGVYLFDEPBGKNAVSRTVINGSTKGLDRKDCBNYITLLERSEGL 140 | |
| Db | 163 | LACGNSQKPSCW-LINNLITOTLGRPLGLAPSPSSGNLYLDDQNTYITINLXLSG 221 | |
| QY | 141 | LACGTNARHPSCNVLVGYTP-LGEGRGYAPSPSPENSJLVLEGEVYSTIRK-OEYNG 198 | |
| Db | 222 | S-HKFRRIAGOVETLSDTAMHRPOFVOATVAKHNSYDKIRIFFEPOENSHSPKOPHT 280 | |
| QY | 199 | KIRFRIRISELITSDTYMONPQFKATIVQDQAYIDKITYFFREDNPKNPEPLN 258 | |
| Db | 281 | VPRVGYCCSDGGESELSVYKWTFLKARLACVDYDTGRIYNELODIFTWQAPENSWE 340 | |
| QY | 259 | VSRVACLRDGDGSELSVSKMNTFLKALVCSDAATNKNFRDLQVFLPDPGQWRD 318 | |
| Db | 341 | TLIYGLFSPWNSAVCVFVKNIDHYFKSKANKNHHKPTRPPOCKMNHQVPETF 400 | |
| QY | 319 | TRYGVGFSPWNTSACVSVYSLGIDIKVFTSSLKGGHSSLPNRPGLCPDOOPITPETF 378 | |
| Db | 401 | QVADRYEVADPYQKNNAMFPIQSKYIYTKLVVREVG-GVFNATIFYLTIKGTIH 459 | |

```
QY      379 QVADRHREVAQRYEPKLTPLFHSKYHQKAVAHMQSHGETH-VLYLTLDGTH 437
      460 IYVREDNSNTALNILEINPFOKPAPIQIILLDNTNLKLYNSEWSEVSPDLDCSVIG 519
      438 KYVEPGEOERHSFNFNIMEIOPFRRAAIQIMSLDAERKRLYSSOWSEVSPDLCEVYG 497
      520 NDCFSCEMSDPPLCTWYNNTC-S-FK-QRVSVETGSGPANTLSEMGDHYAPTVVKAHOVS 576
      498 GGGHGLMSDPIYCGMDQGCISITSSERSVLDOSINPAEPH-KE-CPNPKPDKAPLOKYS 555
      577 IPLLNSYLSCPAVSNHADYFWTKDGFTEKCHVKTKHNDCLLIANSTTATNGITHVCNM 636
      556 LAENRYILSCPMESRHATYSMRHKENVQSGCPHGSPFCILFIEULTAQYGHFCEA 615
      637 KEDS 640
      616 QEGS 619

RESULT      2
ENTRY      148744 #type complete
TITLE      semaphorin A - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
          04-Sep-1998

ACCESSIONS
REFERENCE   148744
#authors    Puschel, A.W.; Adams, R.H.; Betz, H.
#journal    Neuron (1995) 14:941-948
#title      Murine semaphorin D/collapsin is a member of a diverse gene
            family and creates domains inhibitory for axonal extension.
#cross-references MIMD:95267431
#accession  148744
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues   1-748 #label RES
#cross-references EMBL:X85990; NID:g854323; PID:g854324

GENETICS
#gene       sema
CLASSIFICATION #superfamily semaphorin
SUMMARY      #length 748 #molecular-weight 82894 #checksum 9017

Query Match      14.8% Score 737; DB 2; Length 748;
Best Local Similarity 33.7%; Pred. No. 8,75e-131;
Matches 168; Conservative 112; Mismatches 166; Indels 52; Gaps 36;

      112 ECMNFVRLHAYNHTLACRTGAFHTCALMFMATAGTASTGPEKLEDGKGTPTDP 171
      125 DCENYITLLERSEGL-LACGTNARHPSG---WNLVNGI-V-V-P--LGEHMGVAFPEP 174
      172 RHPPVVLVGEELYSGV-TADLMDGRFTTFRSLGQNPSLTEPHDSRWLNEPFVAVEM 230
      175 DENSLVLFEGDEYSTIRKQENYNGK-IPRRRIIRGSELYT-S-DIV-MQNPFIRAT-I 229
      231 PSESNDDDKIYFFFPESAWEAPAMGRMSVSVGOICRNDLGGORSI-VNKKTFPLKAR 289
      230 VHODQAYDKIYFFREDNPDKNPF-APLVNSRVVQOLCRDQGESLSYSKWNPTLKKM 288
      290 LVCSVGVSGDTHFDOLODVFLLS-SR-DROTPFLYAVFSTSSGVQSGAVCYSMNDVR 347
      289 LVCS--DAATNKNFNRLQDVFLPDPGQWRDTRVYGVFSNPNW-Y--SAYCVYSIGDID 343
      348 RAFLGLPHKKEGPTHQWVSYQGRVYPPRQMGCSKTFGFSSTKDFPDDVIOGRNHPLM 407
      344 KYFRTS-SLK-G-----YHSSLPNRPCKLPDQ-OPIP-TEEP-QVAD--R-HPEV 387
      408 YNVLPMGG-R-PLFQVGVGYFTQIADRYVAADGH-XDYLFISTDVGIVYKIVSVK 464
      388 AQREVMGPDKTPLF-H--SKYHQVAVAHMQASHGETFHVLYLTLDGTHIKVE-P- 442
      465 GRPNSEGLLEIELOVEEDSAITTSMQISSKROQLVVASRAVAQIALHRTALGRACAE 524
```

```
QY      443 GEGEHSFAFNIMEIOPFRRAAIQIMSLDAERKRLYSSOWSEVSPDLCEVYGCGCHG 502
      525 CCLARDPYCAMDGSACTRPQPTAKRRFRRODIRNDPSTLSCGSSHSVLEKKVLYVES 584
      503 CLMSRDPYCGMDQGRICISITSSERSYL--QSLNPAEPHKECPNPKAPLOKVSIAENS 560
      585 GSAFLCEPDRSLQAHQW 602
      561 -RYILSCPMESRHATYSW 577

RESULT      3
ENTRY      148748 #type complete
TITLE      semaphorin E - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
          04-Sep-1998

ACCESSIONS
REFERENCE   148748
#authors    Puschel, A.W.; Adams, R.H.; Betz, H.
#journal    Neuron (1995) 14:941-948
#title      Murine semaphorin D/collapsin is a member of a diverse gene
            family and creates domains inhibitory for axonal extension.
#cross-references MIMD:95267431
#accession  148748
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues   1-751 #label RES
#cross-references EMBL:X85994; NID:g854331; PID:g854332

GENETICS
#gene       sema
CLASSIFICATION #superfamily semaphorin
SUMMARY      #length 751 #molecular-weight 85259 #checksum 8961

Query Match      14.6% Score 725; DB 2; Length 751;
Best Local Similarity 33.0%; Pred. No. 4,38e-128;
Matches 148; Conservative 110; Mismatches 149; Indels 41; Gaps 28;

      167 SFNPVNVTYSVINBELFSGMTI-DPMGTDAIFRSLTRKMQRLTDOHNSKWLSEPMFVD 225
      171 PSPSPENSLVFEFGDEYSTIRKQENYNGK-IPRRRIIRGSELYTS--DIV-MQNPFIR 226
      226 AHVIDPDTDPNDAKYFFPKERLTDNNRSTKQIHIMIRICPNDGQORSI-VNKKTFEL 284
      227 ATIVHOD-QAYDKIYFFREDNPDKNPEAPLVNSRVVQOLCRDQGESLSYSKWNPTL 285
      285 KARLYCSYTDDEGPETHDEDELVELI-ETDNP-RTTLVYGIFTSSSVFSGSAVCVYHL 342
      286 KAMLYCS--DA-ATNKNFNRLQDVFLPDPGQWRDTRVYGVFSNPNW-Y--SAYCVYSI 339
      343 SDIQVFNQFPKAKGPNHQLISYQGRIPYPPGTCPGGATFPNNKRTTKDFPDDVYFIR 402
      340 GDIDVYFRTS-SLK-G-----YHSSLPNRPCKLPDQ-OP-I-PTEP-QVADRH 385
      403 NHPLMYNSISPIHRRLPLVIRIGTDYKTKIADVDRNADG-RYHVLFLGTGRGTQKQYV 461
      386 EYAQRYEPMGP-L-KTPLF-H--SKYHQVAVAHMQASHGETFHVLYLTLDGTHIKVE 441
      462 IPTNSASGELLEIELEFKNHVPTTMEISSKQOQLVSSNEGVQSLHRCIHATYAC 521
      442 -PGEDEHSFAFNIMEIOPFRRAAIQIMSLDAERKRLYSSOWSEVSPDLCEVYGCGC 500
      522 ADCCLARDPYCAMDGSHCSREYPTGKRSRQDVHGNPLTQCGRFNLKAVRNAEIVQY 581
      501 HCLMSRDPYCGMDQGRICISITSS--RSVLOSINPAEPHKECP--NPKPKAPLOKVS 556
      582 GVRNNSTFLECAPKSPQASIKWLQDK 609
      557 A-PNSRYILSCPMESRHATYSW-RKEN 582

RESULT      4
ENTRY      G01856 #type complete
```


| TITLE | semaphorin V - human | formal_name | Hom | sapiens | #common_name | man |
|-----------------------|---|--|----------------|--------------|--------------|-----|
| ORGANISM | | | | | | |
| DATE | 21-Dec-1996 | #sequence_revision | 06-Jun-1997 | #text_change | | |
| ACCESSIONS | G01856 | | | | | |
| REFERENCE | G08634 | | | | | |
| #authors | Sekido, Y. | | | | | |
| #submission | submitted to the EMBL Data Library, June 1995 | | | | | |
| #accession | G01856 | | | | | |
| #status | preliminary; translated from GB/EMBL/DBJ | | | | | |
| #molecule_type | mRNA | | | | | |
| #residues | 1-749 | #label | SEK | | | |
| #cross-references | EMBL:U28369; NID:g974283; PID:g974284 | | | | | |
| CLASSIFICATION | #superfamily | semaphorin | | | | |
| SUMMARY | #length 749 | #molecular_weight 83121 | #checksum 2747 | | | |
| Query Match | 14.4%; Score 716; DB 2; Length 749; | | | | | |
| Best Local Similarity | 33.0%; Pred. No. 4,61e-126; | | | | | |
| Matches | 167; Conservative 115; Mismatches 170; Indels 54; Gaps 41 | | | | | |
| Db | 112 | ECMNFVLLHAYNRHLLACGTGAHPICAAVEVGNHAEEDVYLLDPRGIDGKSPYD | 171 | | | |
| Qy | 125 | DCENITTLER-RSEGLACGTNRHPSC-MNLV-N-G-TVYPL-G-EM-NGYAPFS | 173 | | | |
| Db | 172 | PNHRAASVLYGEVLTSGVA-ADIMGRTDITFRSUGORSLTEPHDSRWLKEFVFW | 230 | | | |
| Qy | 174 | PDENSLVLEGGDEVSTLRKQOYNGK-IPRRIRIGSESLT-S-DIV-MONPOFIAT- | 228 | | | |
| Db | 231 | IPESNPDDDKIYFFRETAEAAPALGRLSVRYGQICRNDVGORSI-VNKKTTFLKA | 289 | | | |
| Qy | 229 | IYHQDQAADDKIYFFRRDNPDKNE-AFLNVSRYAQLCRDGGESLSVSKNFTLKA | 287 | | | |
| Db | 230 | RLVCSVPEGDTDFDLDVPLLS-SRDH-RTPLLYAVSTSSIFOGSAVCYISNDV | 347 | | | |
| Qy | 288 | MLVCS--DAATNKNRNRIADYFLLPDPGQMRDTRVYGVFENPMN-Y-SAVCYISLGI | 342 | | | |
| Db | 348 | RRALGFRANHGEGRYHQWYSGRYVRPGMCPKSTGTSTSSKKDPPDDVIOGARNHPL | 407 | | | |
| Qy | 343 | DK--V--F--RSTSS--K--GIHSSLPNDRPGKCLPDD--QPIP-TETP--QVAD--R-HPE | 386 | | | |
| Db | 408 | MYNSSLPTFG-R-PLFLOVGNANTYFTQIAADRVAAADGH-YDVLFIGDVTGLKVISVP | 464 | | | |
| Qy | 387 | VAGRVEPMPGLKTRF-H-SKYNHQKAVHMQASHETHVLYLTDRGTHIKVYE-P | 442 | | | |
| Db | 465 | KGSRPDAEGLLELHVEDSAAVTSQWISSKRHOLYVASSAVAQIALHRCAAGRYCT | 524 | | | |
| Qy | 443 | -GEOHSPAFNIMEIOPRRRAAIAQIOTMSLDERRKLYVSSOMEVSOYPLDICEYGGGCH | 501 | | | |
| Db | 525 | ECCLLARDYICAMDYACITRFQPSAARRRRRDYRNGDPSITCSDGSSRPALLEKRYGVE | 584 | | | |
| Qy | 502 | GCLMSRDRYCGMDQRCISITSSSESYV--OSINPAEPHKECPMPKPPKAPL-QKVSILAP | 558 | | | |
| Db | 585 | GSAFLCEPRLSLOARVEMTFORACV | 610 | | | |
| Qy | 559 | NSRYLSCPMSRRATYSMR-HKENV | 583 | | | |
| RESULT | 5 | | | | | |
| ENTRY | G02173 | #type | complete | | | |
| TITLE | semaphorin III family homolog - human | | | | | |
| ORGANISM | #formal_name Homo sapiens | #common_name | man | | | |
| DATE | 21-Dec-1996 | #sequence_revision | 06-Jun-1997 | #text_change | | |
| ACCESSIONS | G02173 | | | | | |
| REFERENCE | G09275 | | | | | |
| #authors | Naylor, S. | | | | | |
| #submission | submitted to the EMBL Data Library, October 1995 | | | | | |
| #accession | G02173 | | | | | |
| #status | preliminary; translated from GB/EMBL/DBJ | | | | | |
| #molecule_type | mRNA | | | | | |
| #residues | 1-753 | #label | NMY | | | |
| #cross-references | EMBL:U38276; NID:g1061350; PID:g1061351 | | | | | |
| CLASSIFICATION | #superfamily | semaphorin | | | | |

[illegible]

```
QY 133 LERSESG-LIACGTNARHSC-W-NL-VN--GTVVPL--G--EM-RGVAFPSPDENSIVL 181
Db 181 LVDELSGTA-ADFMGRDAIFRTLGHHNPIREQDHSRLNPRIRISAHILIESNPE 239
QY 182 FEGDEVSTIRKOEYNGK-IPFRIRIGESELVTS--DTV-MONPOFIKATIVHQ--DOAY 236
Db 240 DDKIYFFERENALDGEHTGKATHARIGQICNDGFGHSL-VNKTFTFLARLICSPGP 298
QY 237 DDKIYFFEREDNPDKNPEAPLNVSRVAQLCGDGGESSLSVSKWNFTFLAMLVCS--DA 294
Db 299 NGIDTTHDELQDVFLM-NSKDP-KNPIYGVFTTSSNIFKSAVCMSMTDVRVFLGPY 356
QY 295 -ATNKNFNRLQDVFLLPDPGQWMDTRVGVFSNPMN-Y--SAVCYSLGIDIKVFRTS- 349
Db 357 AHDGPNYQWVYOGRVYPRPGTSPKSTGCGEFTSDLDDEVTITPARKSHPMNYPVFP 416
QY 350 SLK-G-YH--S--S-LPNRPGKCLPQO-OP1-PTETF--QV-A-DR-HPEVAORVEPM 394
Db 417 NS-RPIWIKTDVYQFTQIIVDRVDAEDGO-YDVMFICTGTGTVLKVYSIPKFTWHELE 474
QY 395 GPLKTPLEHFK--YHQQKAVHMQASHGETFHVLYLTTRGTIHKVYE-PGEDEHSPAF 451
Db 475 VLEEMTVFREPVIYISAMKISTQOOLYISATGVSQOLPLHRCVDYKACACCECLARDPY 534
QY 452 NIME-IQPFRAAIIQMSIDAERRKLYVSSQWESQVPLDCEVYGGCGHCLMSRDY 510
Db 535 CAMDGSSCSRYFTAKRRTRRDIIRNGDPLTHGSDLDHNDHNPSCQGLEEIIYGVENSST 594
QY 511 CGMDQGRCSITYSSE-RSVL-QSINPAEPHKECPN-PRPKAP--LQ-KVSLA-BNSRY 562
Db 595 FLECSKPSQALVY-WQFQKQNDHKE 621
QY 563 YLSC-PMESRHATYSWR-HKENVEQSC 588

RESULT 7
ENTRY 148747 #type complete
TITLE semaphorin D - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Sep-1998
ACCESSIONS 148747
REFERENCE 148744
#authors Puschel, A.W.; Adams, R.H.; Betz, H.
#journal Neuron (1995) 14:941-948
#title Murine Semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension.
#cross-references MUID:95267431
#accession 148747
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-772 ##label RES
#cross-references EMBL:X65993; NID:g854329; PID:g854330
GENETICS
#gene SemD
CLASSIFICATION #superfamily semaphorin
SUMMARY #length 772 #molecular_weight 88710 #checksum 1776

Query Match 13.5%; Score 669; DB 2; Length 772;
Best Local Similarity 30.2%; Pred. No. 1,56e-115;
Matches 170; Conservative 143; Mismatches 182; Indels 67; Gaps 51;

Db 58 HTFLDE-ERSFLYGADHIFSNLVNIKDFQITVPIVSY-TRDECKMAGKDLKCA 115
QY 75 HTVLFHEGSSSVWVGKGYLDFPEGKN-AS-VRTVIGSTGSC--LDK--RDCE 127
Db 116 NFIKYLEAYNOTHLYACGTGAFHPICTYLEVGHHPEDNIFKLDHFNFGKSPDYDKL 175
QY 128 NYITLLEHRSRG-LIACGTNARHPSCNVL-V-N--GTVVPL--G--EM-RGIAPFSPDE 176
Db 176 LTASLLIGELYSSTA-ADFMGRDAIFRTLGHHNPIRTEQDHSRLNDFPFIASHLPIE 234
QY 177 NSLVFEGDEVSTIRKOEYNGK-IPFRIRIGESELVTS--DTV-MONPOFIKATIVHQ 232
```

```
Db 235 SNPDEDDKYFFERENALIGEHSGKATHARIGQICNDGFGHSL-VNKTFTFLARLIC 293
QY 233 -DOAYDKIYFFEREDNPDKNPEAPLNVSRVAQLCGDGGESSLSVSKWNFTFLAMLV 291
Db 294 SVYSPGIDITHPELDQVFLM-NSKDP-KNPIYGVFTTSSNIFKSAVCMSMTDVRV 351
QY 292 S--DA-ATNKNFNRLQDVFLLPDPGQWMDTRVGVFSNPMN-Y--SAVCYSLGIDIKV 345
Db 352 FLGPNARBDPNQWVYOGRVYPRPGTSPKSTGCGEFTSDLDDEVTITPARKSHPMN 411
QY 346 FRIS-SLK-G-YH--S--S-LPNRPGKCLPQO-OP1-PTETF--QV-A-DR-HPEVAQ 389
Db 412 PVFPINN-RPIWIKTDVYQFTQIIVDRVDAEDGO-YDVMFICTGTGTVLKVYSIPKFTW 469
QY 390 RVEPMGPLKTPLEHFS--KYHQKAVHMQASHGETFHVLYLTTRGTIHKVYE-PGEDE 446
Db 470 HDLEVLDEEMTVFREPVIYISAMELSTKQOOLYISATGVAQLPLHRCDIYKACACCECL 529
QY 447 HSFAPNIME-IQPFRAAIIQMSIDAERRKLYVSSQWESQVPLDCEVYGGCGHCLM 505
Db 530 ARDPYCAMDGSSCSRYFTAKRRTRRDIIRNGDPLTHGSDLDHNDHNPSCQGLEEIIYGV 589
QY 506 SRDPYCGMDQGRCSITYSSE-RSVL-QSINPAEPHKECPN-PRPKAP--LQ-KVSLA- 557
Db 590 ENSSTFLECSKPSQALVY-WQ 610
QY 558 PMSRYLSC-PMESRHATYSWR 578

RESULT 8
ENTRY 158169 #type fragment
TITLE semaphorin III - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 04-Sep-1998
ACCESSIONS 158169
REFERENCE 158169
#authors Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.;
#journal Tessler-Lavigne, M.; Goodman, C.S.; Kolodkin, A.L.
#title Semaphorin III can function as a selective chemorepellent to pattern sensory projections in the spinal cord.
#cross-references MUID:95267432
#accession 158169
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-666 ##label RES
#cross-references GB:L40484; NID:g703189; PID:g703190
GENETICS
#gene SemIII
CLASSIFICATION #superfamily semaphorin
SUMMARY #length 666 #checksum 9654

Query Match 13.4%; Score 667; DB 2; Length 666;
Best Local Similarity 30.1%; Pred. No. 4.37e-115;
Matches 150; Conservative 130; Mismatches 168; Indels 50; Gaps 39;

Db 60 RGKSPYDKLTASLLIGELYSSTA-ANFMGRDAIFRTLGHHNPIRTEQDHSRLNDP 118
QY 167 RGVAFPSPDENSIVLFEQDEVSTIRKOEYNGK-IPFRIRIGESELVTS--DTV-MONP 222
Db 119 RFIHAHLIPESDNEDDKYFFERENALDGEHTGKATHARIGQICNDGFGHSL-VNKN 177
QY 223 QFIKATIVHQ--DOAYDKIYFFEREDNPDKNPEAPLNVSRVAQLCGDGGESSLSVSKW 261
Db 178 TTFIKALICVSPGPNIDITHPELDQVFLM-NSKDP-KNPIYGVFTTSSNIFKSAVC 235
QY 282 NTFIKAMLVCS--DA-ATNKNFNRLQDVFLLPDPGQWMDTRVGVFSNPMN-Y--SAVC 335
Db 236 MYSMDVRRVLYGVAHARBDPNQWVYOGRVYPRPGTSPKSTGCGEFTSDLDDEVTIT 295
QY 336 VSLGIDIKV---F--RTSS-LK--GYHSSLPNRPGKCLPQO-OP1-PTETF--QV-A 381
```

| | | | |
|--|-------|--|--------------------------------|
| Db | 236 | FASHPAMNYVPEPLNN-RPIMITDVYQSTQIVNVRVADSDGO-XDVMEIGDVGTVL | 353 |
| Oy | 382 | -DR-HPEVAQVREBPGLKPLTFHS-KYHQAQVAMHMOASHETFNHLYLTDRCTIH | 437 |
| Db | 354 | KVSVPEKETMWDLEVLLEEMTVREPTITAMELSTRKQOOLYIGTAGVACLPHACDI | 413 |
| Oy | 438 | KVVE-PGEOHSAFNINME-IQPFRRAAIQTMSLDAERKRLYVSSQMEVSOVPDLCEV | 495 |
| Db | 414 | YKACAEECLARDPYCAWDGSSCSRYEFTPAKRTRRODIRNGDPLTHCSDILOHHNHGP | 473 |
| Oy | 496 | YGGCGHCLMSRDPYCGWDGRCISYSSSE-RSVL-QSINPAEPHKRCPPN-KRPDK--AP | 550 |
| Db | 414 | SLEENITIGVENSSFFELCSKSORALYYWOPORNRNDRKEIKMGHIIITDEGLLRS | 533 |
| Oy | 551 | -LQ-KVSLA-PSNRYLLSC-PMESHRYAT-SW-RHKENVEGSCBPG-HQ-SPNCILFIEN | 602 |
| Db | 534 | LQKRDGNYLCHAVEHGF | 551 |
| Oy | 603 | LTAQOYGHYFCEADGEGST | 620 |
| RESULT | 9 | | |
| ENTRY | | D49423 | #type complete |
| TITLE | | Semaphorin III precursor - human | |
| ORGANISM | | Homosapiens | #common_name man |
| DATE | | 06-Jan-1995 | #sequence_revision 06-Jan-1995 |
| ACCESSIONS | | D49423 | |
| REFERENCE | | D49423 | |
| authors | | Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S. | |
| #journal | | Cell (1993) 75:1389-1399 | |
| #title | | The Semaphorin genes encode a family of transmembrane and secreted growth cone guidance molecules. | |
| #accession | | D49423 | |
| #status | | preliminary; nucleic acid sequence not shown | |
| #molecule_type | | mRNA | |
| #residues | | 1-771 | #label KOL |
| GENETICS | | #cross-references GB:126081; NID:9799328; PID:9436560 | |
| #gene | | GDB:SEMA1 | |
| CLASSIFICATION | | #cross-references GDB:283448 | |
| SUMMARY | | #superfamily semaphorin | |
| | | #length 771 | #molecular_weight 88889 |
| | | | #checksum 6249 |
| Query Match | 13.3% | Score 661; DB 2; length 771; | |
| Best Local Similarity | 30.3% | Pred. No. 9.55e-114; | |
| Matches 151; Conservative 126; Mismatches 174; Indels 47; Gaps 37; | | | |
| Db | 166 | RKSPYDPKLLTASLLIDGELYSGT-ADFKGRFAFRLGHHNPIRTEDHDSRWLNDP | 224 |
| Oy | 167 | RGVAFSPDENSIVLFESEDEVYSTRKEQYNGK-IPFRFRRGSESLYTS-DIV-MQN | 222 |
| Db | 225 | KFISAHLISESDNPBDKVFYFFPENALIDGSHSKATHARIGQICKNDFGCHSL-VNKK | 283 |
| Oy | 223 | QFIKATIVHQ-DOAYVDKDIYFFREDNDKMPKAPLNVSRVAQACRCDGEGESLSYSKW | 281 |
| Db | 284 | TFPLKARLISVPGNGIDTHEDLQVYELM-NKND-KNRYVYGVYTTSSNIFKSGAY | 341 |
| Oy | 282 | NFLPKAMLYCS-DA-ATNKNFNRLQDFLLPDRSGMGRDTRVYGVSNPN-Y-SAVC | 335 |
| Db | 342 | MYSMSDVARVYLGFAVHRDGPYQVMPVQGVPRYPREGTSPSKFFGFSSTKDLPDVIT | 401 |
| Oy | 336 | VYSLGDDIDKVERTS-SLK-G-YH-S-S-LPVRPRGKCLPDQ-QPI-PTETP-QV-A | 381 |
| Db | 402 | FARSHPAMNYVPEPLNNRPIYIKTDVNYQTOQIVNVRVADSDGO-YVMFIGDVGTVL | 460 |
| Oy | 382 | -DR-HPEVAQVREBPGLKPLTFHS-KYHQAQVAMHMOASHETFNHLYLTDRCTIH | 438 |
| Db | 461 | VVSPKPEWYDLEVLLEEMTVREPTITAMELSTRKQOOLYIGTAGVACLPHACDI | 520 |
| Oy | 439 | VVE-PGEOHSAFNINME-IQPFRRAAIQTMSLDAERKRLYVSSQMEVSOVPDLCEV | 496 |

| | | | |
|----|-----|--|-----|
| Db | 521 | GKAACECLLRDPICAMDGASGRYPITAKRRTRRQDINDGNPILFICSPDLHNDHNGSP | 580 |
| Oy | 497 | GGGCHGLMRSRDPICMGDGRCLSIYSSE-RSLV-QSIPAEPRHKECPKPKDK---- | 550 |
| Db | 581 | ERIITYGVENSSTFLCSPKSPQALVYWOFRNREERKEEIRVDHIIRTDGGLLSRLQ | 640 |
| Oy | 551 | LQVSLA-PSNRYLSC-PRESHRATY-SW-RKKEV-EGCEPGHQ-SPNCILFLENLT | 604 |
| Db | 641 | OKDSGNTLCHAVHEGFIO | 658 |
| Oy | 605 | AQQGYHFCFAOEGSEYER | 622 |

| | |
|--------------------|--|
| RESULT | 10 |
| ENTRY | I48746 |
| TITLE | semaphorin C - mouse (fragment) |
| ORGANISM | #formal_name Mus musculus #common_name mouse |
| DATE | 02-Jul-1995 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997 |
| ACCESSIONS | I48746 |
| REFERENCE | I48744 |
| #authors | Puschel, A.W.; Adams, R.H.; Betz, H. |
| #journal | Neuron (1995) 14:941-948 |
| #title | Mutine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension |
| #cross-references | MIMD:95267431 |
| #accession | I48746 |
| #status | preliminary; translated from GB/EMBL/DBJ |
| #molecule_type | mRNA |
| ##residues | 1-782 ##label RES |
| ##cross-references | EMBL:X85992; NID:9854327; PID:9854328 |
| GENETICS | |
| #gene | semc |
| SUMMARY | #length 782 #checksum 1571 |

| | | | |
|----|-----|--|-----|
| Db | 74 | KRDQNTIKILLPLNSHLTCTGAFFSPICAVIHIASTFLADQAGNYILEDGRHCPF | 133 |
| Oy | 123 | KRDQNTIT-LLERRSEGLAGCTINRHPSC-W-N-----LV-NGIV-VPLGEMGVIAPF | 172 |
| Db | 134 | DPNFKSTALVYDGLHYGTVS-SFOGNDPAISRQSSRPTKESSLNMLDPAFAVSAATS | 192 |
| Oy | 173 | SPDENSELVLEEGDEVYSTIRKQYKNGKIRFRIRRESLHYSDIV--MQNPQIFKATIV | 230 |
| Db | 193 | PESIGSPIGDDKIYFFSEFGQEEFFENTI-VSNVAHVCKGDEGGERVLD-QRWISFL | 250 |
| Oy | 231 | HQDQ-A-Y--DKRIYEFFREDNDP-KNPAPLAVSVHVAOLCGDGGESSLSVSKMNTFL | 285 |
| Db | 251 | KAOILLCGRPDGPPENVLQDVFYTLNPNQ-DNRKLTISIVFTSQMHRGRTBESALCYETM | 309 |
| Oy | 286 | KAMLVYCDATATNKNFNKLQDVFLL-DBPFGQWMDTRFVYVSFSPNPN--Y--SAVCYSL | 339 |
| Db | 310 | NDVOKAPADGLYKKVKNRETOQYETETHQVTPRPGACITNSAREKINSLSQLPDVLNFL | 369 |
| Oy | 340 | GDIDKVF-----RTSS-LKG-Y-HSS-LNNPRPGKLPD--QQPIPTETFGVADNHPEVA | 388 |
| Db | 370 | KDHELMADGVRSRILLQPPARAYQAVAHVPGDLS-TYDVLFCTGDGRLLKAVATLSR | 428 |
| Oy | 389 | QRVEPM-GPLKTPLE--HSKYHYQKAVAHVPMQASHGETPHVLYLTTRDRIKHYVEPEQ | 445 |
| Db | 429 | VH---I-IEELQIFPQOQVPYQNLDSHGGLYASHSHSVQVQPVANCSLYPT-IGCDLL | 483 |
| Oy | 446 | EHSFAFMIMELQFFRRRAALQIWSLDAERKRKLVSQWQEVSOVPLDCEYVGCGHGLM | 505 |
| Db | 484 | ARDEYCAWTSACRLASIYOPD | 505 |
| Oy | 506 | SRDPYCGMDGRC-I-SIYSSE | 525 |

| | |
|--------|----|
| RESULT | 11 |
|--------|----|

| | | |
|-----------------------|---|---|
| ENTRY | 656498 | #type complete |
| TITLE | M-sema F protein precursor - mouse | |
| ORGANISM | #formal_name Mus musculus #common_name house mouse | |
| DATE | 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997 | |
| ACCESSIONS | 566498 | |
| REFERENCE | 566498 | |
| #authors | Inagaki, S.; Furuyama, T.; Iwashashi, Y. | |
| #journal | FEBS Lett. (1995) 370:269-272 | |
| #title | Identification of a member of mouse semaphorin family. | |
| #cross-references | identical MUID:95385809 | |
| #accession | 566498 | |
| ##status | preliminary | |
| ##molecule_type | mRNA | |
| ##residues | 1-834 #label INA | |
| ##cross-references | EMBL:579463; NID:g1110598; PID:g1110599 | |
| FEATURE | | |
| 1-21 | #domain signal sequence #status predicted #label SIG | |
| 22-834 | #product M-sema F protein #status predicted #label MAT | |
| SUMMARY | #length 834 #molecular_weight 92556 #checksum 7189 | |
| Query Match | 10.5%; Score 522; DB 2; Length 834; | |
| Best Local Similarity | 28.5%; Pred. No. 5,15e-83; | |
| Matches | 142; Conservative 118; Mismatches 19; Indels 49; Gaps 36 | |
| Db | 102 | KGKS-NQTECFNFIFFLDPYNSHLVYCCTAFQPKCTIYIMLTLLDRAEFEDGKGP 160 |
| Qy | 117 | KGSCDKRRCENYITLLER-RSEGLLAGGTNRHPSC-W-NLVNGTV--VPLGHRGYAP 171 |
| Db | 161 | YDPAGHGTGLVDGELGYSLTLN-NELGPEPVLTIRMGTGHTSKTEYLAFWLNEPHTVSA 219 |
| Qy | 172 | FSPDENSELVLEEGDEVYSTIKQETNGKIPFRFRRIRGESELYTSD--TV-MQNPQFIAT 228 |
| Db | 220 | FVPEVSGETGDDDKIYFFFSERAVEYDCSEQVARYARYCKGDMGARTIQ-KKATTF 278 |
| Qy | 229 | IVHOD-QAY--DCKIYFFFRDNDPKNEAPLANSRAQOLCRGOGGESSLSYSKMTF 284 |
| Db | 279 | LKARLYGAPDMKYVFNQIKAVHTLRGAS--WHNTTFPGVFOARWMDLSAVCEYOLEQ 336 |
| Qy | 285 | LKAMLVGSDAATNKNFNRLQDVFLLPDSGQWRDTRVYGVFSPW---NYSACVYSLGD 341 |
| Db | 337 | IQQVEGEGYKKESEDQAQWARTDVPSPRQSCINMHRNDGYTSSLEDPNTLNTFKK 396 |
| Qy | 342 | IDKVF---RT-S--SLK-G-YHSSLPNPRPKCLPD--QQPIP-TETFOVADRPE-VA- 388 |
| Db | 397 | HPLMEDQYKPRGLRPLKKNNTFNFNVAADRYGADGATYVLFCTGDGWLKAVSGP 456 |
| Qy | 389 | QR-VE-PNGP-LKPTLFLFSKY-HYQKAVAVHNRQASHGETFHVLYLTTRGTLTHKAYEGE 444 |
| Db | 457 | WIH---M-VEELQYFQDEP-VESTLVLSQSKYVLFAGSRSSQVLQSLADCTKYRF-CYDCV 510 |
| Qy | 445 | OEHSPAFNIMEIQPRRAAIQIOTMSLDABRRRLVYSSQMEVSYQVPLDICEYVGGGCHCL 504 |
| Db | 511 | LARDPYCAMNNTNSCAVTTSGRGSSELYQVANYADTSMKQNGIKKRVISPIKNIYVS 570 |
| Qy | 505 | MSRDPYGCMD-Q-GRCISITYSSR-SVL-QSINPAEPKPECPNPKPDKA-PLOK-VSLAP 558 |
| Db | 571 | GTDLVLPCHLSSNLNAHNM 589 |
| Qy | 559 | NSRYLSCPMESRATYSM 577 |
| RESULT | 12 | |
| ENTRY | 148745 | #type complete |
| TITLE | semaphorin B - mouse | |
| ORGANISM | #formal_name Mus musculus #common_name house mouse | |
| DATE | 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 28-Feb-1997 | |
| ACCESSIONS | 148745 | |
| REFERENCE | 148744 | |
| #authors | Puschel, A.W.; Adams, R.H.; Betz, H. | |
| #journal | Neuron (1995) 14:941-948 | |
| #title | Murine semaphorin D/collapsin is a member of a diverse gene | |

| | | | | | | |
|-----------------------|--|--------------------|--|--------------|---|---------|
| | family and creates domains inhibitory for axonal extension | | | | | |
| #cross-references | MUID:95267431 | | | | | |
| #accession | I48745 | | | | | |
| #status | preliminary; translated from GB/EMBL/DDBJ | | | | | |
| #molecule_type | mRNA | | | | | |
| #residues | 1-760 #label RES | | | | | |
| #cross-references | EMBL:X85991; NID:g854325; PID:g854326 | | | | | |
| GENETICS | | | | | | |
| #gene | smb | | | | | |
| SUMMARY | length 760 #molecular-weight 83458 #checksum 2188 | | | | | |
| Query Match: | 9.7%; Score 480; DB 2; Length 760; | | | | | |
| Best Local Similarity | 29.2%; Pred. No. 7,12e-74; | | | | | |
| Matches | 157; Conservative 121; Mismatches 203; Indels 57; Gaps 411 | | | | | |
| D | B | 123 | QCFNTRIVLVSINATHLYACGTFAPFSPACTIELODSLILPILIDIKYWDGKGSPLTFT | 182 | : | : : : : |
| Q | Y | 125 | DENENTTLDER-RSGGLLACGTNAHPSC-N-NLVNGTVVP-LGE-M-RGYAPFSP-D | 175 | : | : : : : |
| D | B | 183 | STQAALVVG-MLYSGT-NNNFSGSEPIIMRLTGSHPVAKTDIFLRMLH-ADASPVAIPS | 239 | : | : : : : |
| Q | Y | 176 | ENSLVLFEDDEYSTRKQETNGKIPRRIRRGESELTSJTVQNQPDKATIVHDDA | 235 | : | : : : : |
| D | B | 240 | -TQVVYFEFEETASEDFEELIYSRVAOVCKNDVGGELIQ-KKWTFLKAQLLCAPG | 297 | : | : : : : |
| Q | Y | 236 | YDDKIYEFREDNPDKNEAPLVNSRVQAOLCRGDGGSSLSVKMTFLKAMLVCSDA | 295 | : | : : : : |
| D | B | 298 | -QLPNITRHAVLLPADS-P-SVSRITAVFTSQNOVGTSASSAVCAFSLDDIERVEFKY | 354 | : | : : : : |
| Q | Y | 296 | TKNKNRNRIADVFLLPDPGOWRDTRVGVGFSPNPNY-----SAVCVYSTLGIDIKVFR-T | 348 | : | : : : : |
| D | B | 355 | KELNKETRMTYTGRSEYSPRGSCSMGPSDDKALT-FMKD-HFLMEHVYGT-PL--L | 408 | : | : : : : |
| Q | Y | 349 | SSL-K-G-----YHSLEPNRPAGKCLPRQQPIPETQTQVADRHEVGAQRPEMPGLTPL | 401 | : | : : : : |
| D | B | 409 | VKSGETRFLAVESARGLDGSSHVMYLIGTSGPLRAVVP-OOSS-AVLVEEIQLSPD | 465 | : | : : : : |
| Q | Y | 402 | FHSKHYYQKAVHVRMQASHGETFMHYLYLTDRGTIHKKVEREGEHSAFNIMEIOFRR | 461 | : | : : : : |
| D | B | 466 | SEPPVNLDLARQGAVPFGFSGGILTRVPNRANCYYES-CVDCLARPHCAMDEPSALCS | 524 | : | : : : : |
| Q | Y | 462 | AAAIIDMSLDERRLRYSSOMEWSOVPLDICFYGGCGHCCLMSRDPYGMD-QGR-CI | 519 | : | : : : : |
| D | B | 525 | LLSGSTKWKMDMENGANEWCRTKGPMARSPPROSPOLIKEYVTYNSTIELRCPLISA | 584 | : | : : : : |
| Q | Y | 520 | SYSSERBVLOSINAEHKRCP-NP-K-PDK-AP--LQKVSLA-PNSKYILSCPHESR | 571 | : | : : : : |
| D | B | 585 | LASYWSHGKRAKISEASATYVNGSL-LLLPDGVGGLY-Q-CVATENGXYTPVSYW | 638 | : | : : : : |
| Q | Y | 572 | HATYSWRH-KENVQSCPEGHQSPNCILFIENLTAQOYGHVFCEAQGSY-FREAQM | 627 | : | : : : : |
| RESULT | 13 | | | | | |
| ENTRY | E42521 | #type complete | | | | |
| TITLE | A39r protein - vaccinia virus (strain Copenhagen) | | | | | |
| ORGANISM | #formal_name Vaccinia virus | | | | | |
| #note | host Homo sapiens (man) | | | | | |
| DATE | 09-Nov-1990 | #sequence_revision | 03-Nov-1990 | #text_change | | |
| ACCSSIONS | E42521 | | | | | |
| REFERENCE | A33172 | | | | | |
| #authors | Johnson, G.P. | | | | | |
| #submission | submitted to Genbank, June 1990 | | | | | |
| #accession | E42521 | | | | | |
| ##status | preliminary | | | | | |
| ##molecule_type | DNA | | | | | |
| ##residues | 1-403 #label JOH | | | | | |
| SUMMARY | length 403 #molecular-weight 45741 #checksum 8167 | | | | | |
| Query Match | 9.5%; Score 470; DB 2; Length 403; | | | | | |
| Best Local Similarity | 33.5%; Pred. No. 1.04e-71; | | | | | |
| Matches | 86; Conservative 59; Mismatches 95; Indels 17; Gaps 16; | | | | | |

| | | | |
|-----------------------|-----|--|---|
| Db | 77 | LVCGTNNNGNPPCKW - IOSDDPKHGRGRYAYQNSKVTIISHNCC - VLSOLINISK - GG - I | 132 |
| Qy | 141 | LACGTNAHPSHCMLVNCVTVPFLGEMRGYAFSPDENSESLVFEDEYASTIRKQETNGKI | 200 |
| Db | 133 | KRMRFDPGCGDYDTADNVIPKDG - LRGAFVDRKDTYD - KVTILFDTITGSKR - I - VK | 187 |
| Qy | 201 | PFRFRIRKE - S - ELTYSUTVWQNOQFIKATIVHODQAVDKITVFFREDNDKKNPEAPLN | 258 |
| Db | 188 | IPYIAOMCLANDEGGPSSLSHSRWSTFLKVELEC - DID - GRSY - R - Q - IHSRTIKTD - ND | 241 |
| Qy | 259 | VSRAVAOLCRGQGGSSLSVSKWMTFLKAMLVCSDAATNKNFNRLQDVFLLPDPSCGQWRD | 318 |
| Db | 242 | TILVYFEDSPYSKSLCTIYSMNTIKQSTSTKLEGYTKOLPSPAPGICLPAGKVVPHTF | 301 |
| Qy | 319 | TRVGVESNPWNYSVAVCYISLGIDIKVFRISLKGHYHSLPNRPGRKCLPDPQOPPIETEF | 378 |
| Db | 302 | EVIEKYNVLDDIIRPLS | 318 |
| Qy | 379 | QVADRHREVAORVERPMG | 395 |
| RESULT | 14 | | |
| ENTRY | | S29921 | #type complete |
| TITLE | | hypothetical protein 15 - vaccinia virus | |
| ORGANISM | | #formal name vaccinia virus | |
| DATE | | 20-Feb-1995 | #sequence_revision 20-Feb-1995 |
| | | 09-Sep-1997 | #text_change |
| ACCESSIONS | | S29921 | |
| | | S29907 | |
| #authors | | Amegadzie, B. Y. | |
| #submission | | submitted to the EMBL Data Library, January 1991 | |
| #accession | | S29921 | |
| #status | | preliminary | |
| ##molecule | | DNA | |
| ##residues | | 1-441 | #label AME |
| ##cross-references | | EMBL:X57318; NID:962239; PID:962254 | |
| SUMMARY | | #length 441 | #molecular_weight 50185 |
| | | | #checksum 6034 |
| Query Match | | 9.3%; | Score 463; DB 2; Length 441; |
| Best Local Similarity | | 33.9%; | Pred. No. 3.37e-70; |
| Matches | | 87; | Conservative 57; Mismatches 96; Indels 17; Gaps 16; |
| Db | 115 | LVCGTNNNGNPPCKW - IOSDDPKHGRGRYAYQNSKVTIISHNCC - VLSOLINISK - EG - I | 170 |
| Qy | 141 | LACGTNAHPSHCMLVNCVTVPFLGEMRGYAFSPDENSESLVFEDEYASTIRKQETNGKI | 200 |
| Db | 171 | KRMRFDPGCGDYDTADNVIPKDG - LRGAFVDRKDTYD - KVTILFDTITGSKR - I - VK | 225 |
| Qy | 201 | PFRFRIRKE - S - ELTYSUTVWQNOQFIKATIVHODQAVDKITVFFREDNDKKNPEAPLN | 258 |
| Db | 226 | IPYIAOMCLANDEGGPSSLSHSRWSTFLKVELEC - DID - GRSY - R - Q - IHSRTIKTD - ND | 279 |
| Qy | 259 | VSRAVAOLCRGQGGSSLSVSKWMTFLKAMLVCSDAATNKNFNRLQDVFLLPDPSCGQWRD | 318 |
| Db | 280 | TILVYFEDSPYSKSLCTIYSMNTIKQSTSTKLEGYTKOLPSPAPGICLPAGKVVPHTF | 339 |
| Qy | 319 | TRVGVESNPWNYSVAVCYISLGIDIKVFRISLKGHYHSLPNRPGRKCLPDPQOPPIETEF | 378 |
| Db | 340 | EVIEKYNVLDDIIRPLS | 356 |
| Qy | 379 | QVADRHREVAORVERPMG | 395 |
| RESULT | 15 | | |
| ENTRY | | JH0798 | #type complete |
| TITLE | | fascioli IV precursor - American bird grasshopper | |
| ORGANISM | | #formal name Schistocerca americana | #common name American bird grasshopper |
| DATE | | 30-Sep-1993 | #sequence_revision 30-Sep-1993 |
| | | 09-Sep-1997 | #text_change |
| ACCESSIONS | | JH0798 | |
| REFERENCE | | Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.; Patel, N.H.; | |
| #authors | | | |

```

#journal      Admon, A.; Bentley D.; Goodman, C.S.
#title       Fascilin IV: Sequence, expression, and function during
#accession   JH0798      growth cone guidance in the grasshopper embryo.
                  ##molecule 1-730
                  ##residues 1-730
                  ##cross-references GB:L00709; NID:g160844; PID:g160845
                  ##experimental source embryo
COMMENT      This protein plays a role in growth cone guidance in the developing
                  central nervous system.
KEYWORDS     glycoprotein; transmembrane protein
FEATURE
1-22        #domain signal sequence #status predicted #label S1G\
23-627      #product fascilin IV #status predicted #label M1\
628-652     #domain extracellular #status predicted #label E1\
653-730     #domain transmembrane #status predicted #label TM\
44,71,163,267,360, #domain intracellular #status predicted #label I1\
539         #binding_site carbohydrate (Asn) (covalent) #status
                  predicted
SUMMARY      #length 730 #molecular-weight 81214 #checksum 5881

Query Match      9.3%: Score 461; DB 2; Length 730;
Best Local Similarity 28.0%: Pred. No. 9,11e-70;
Matches 143; Conservative 130; Mismatches 182; Indels 56; Gaps 46;

Db 32 QGEGERYQRFPLGNESEKDHFKLEKKNHSLGAGNIYVNIISLRLDTFTFEQRIEMHSSG 91
Oy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   61 HVGQDRND -FGQTEPHIVLFH--EPSSSVWVGKGYKLPDPFGKN-ASVRIY-NI-G 114
   92 AHRELCLYKGSDDCCNTIRVLAKIDDRVLICGTNAVKPLCRHYALKDDGYVEKEYE 151
Oy : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   115 STKGSC-L-DKR--DDENYITLERSSE6-LIACGTNARHPSQW-N-LVNGT-VVPLG-E 165
Oy 152 GRGICPPDPDHNSTAIYSGOLYARV-ADFGSDP-LI-YRG-P-LRTERSGL-K-OLN 204
   166 MRGTAIPSPDENSILVLEGDEVIYSTIRKOEYKGLPRFRRIIGSESLTSDTWQNPQFI 225
Oy 205 APNEFV-NTMEYNDPIFFEFRTAVEXINGCAIY-SRAYVCKHDKGSPHOGF-DRWTSF 261
   226 KATIVHQDAQYDKIYFFREDNP-D-KNEADPLNSRAVALCRGDGESSLSYKNTFE 284
Oy Db 262 LKSLKNSVPGDYPTFFNELQSTDIIEBNGGQ-VKLIYGVFTTPNLSIGSAYCAF5 320
   285 LKALWYCSDAATK-NFNRLDVF-LLPDP-SGQMRDTRVYGVFSPWN-Y-SAVCYA5 338
Oy Db 321 MKSLIEGDFPFEKQETAMSNMVLAVPSLVPEPFGOCQNSRTPDVSANFVSHITMD 380
   339 L-G--D-IDVYFT-SLKG-YHS--SL--PAPRGRKCLPDQOPIPTFTFOVADRHPEVA 388
Oy Db 381 EAVPAFT-RPIIRISLOYRFTKIAVDQOVRTPGKAVDVLIGTDGKVIKALNSAF5 439
   389 QVPEPMPLTPTFFH-S-KYHOKYAV-HRMDSHGTEFHYLYLTDDGITHKVEP-- 443
Oy Db 440 DSSPTDVSVTEELQVLPBGVPRKYLNVYARMGDGSKLVVSDDEILIKIHRGSKIT 499
   444 EQHSIFNFNME-IQFFRAAAIQI---SLAEKRKLIVSSQWVEVSOVPLDLCEVIG-G 498
Oy Db 500 NCRECVSLDPPYCAMDNVELCTAVGSPDMS 530
   499 GCHGLMSRDPYCGMDQ-G-RCISIVSSERS 527

Search completed: Tue Aug 3 16:28:41 1999
Job time : 110 secs.

```

This Page Blank (uspto)

M P E S R E F
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 3 16:32:43 1999; Maspar time 8.46 Seconds
799.386 Million cell updates/sec
Tabular output not generated.

Title: >US-09-240-410-2
Description: (1-666) from US09240410.pep
Perfect Score: 4968
Sequence: 1 MTPPPGRAAPSAPRARVPG.....LAASIMLGVLPTLTGLVH 666

Scoring table:
PAM 150
Gap 11

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCR9_COMB 4:Backfiles1

Statistics: Mean 35.201; Variance 163.168; scale 0.216

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|------------------------------------|-----------|
| 1 | 661 | 13.3 | 771 | 3 | PCT-US94-1 Sequence 54, Applicati | 9.28e-47 |
| 2 | 661 | 13.3 | 771 | 2 | US-08-835- Sequence 54, Applicati | 9.28e-47 |
| 3 | 661 | 13.3 | 771 | 1 | US-08-121- Sequence 54, Applicati | 9.28e-47 |
| 4 | 620 | 12.5 | 477 | 1 | US-08-136- Sequence 2, Applicatio | 4.24e-43 |
| 5 | 463 | 9.3 | 441 | 3 | PCT-US94-1 Sequence 56, Applicati | 3.22e-29 |
| 6 | 463 | 9.3 | 441 | 2 | US-08-835- Sequence 56, Applicati | 3.22e-29 |
| 7 | 463 | 9.3 | 441 | 1 | US-08-121- Sequence 56, Applicati | 3.22e-29 |
| 8 | 459 | 9.2 | 730 | 1 | US-08-835- Sequence 58, Applicati | 7.22e-29 |
| 9 | 459 | 9.2 | 730 | 3 | PCT-US94-1 Sequence 58, Applicati | 7.22e-29 |
| 10 | 459 | 9.2 | 730 | 1 | US-08-121- Sequence 58, Applicati | 7.22e-29 |
| 11 | 378 | 7.6 | 712 | 1 | US-08-121- Sequence 64, Applicati | 7.91e-22 |
| 12 | 378 | 7.6 | 712 | 2 | US-08-835- Sequence 64, Applicati | 7.91e-22 |
| 13 | 378 | 7.6 | 712 | 3 | PCT-US94-1 Sequence 64, Applicati | 7.91e-22 |
| 14 | 367 | 7.4 | 650 | 1 | US-08-121- Sequence 60, Applicati | 7.00e-21 |
| 15 | 367 | 7.4 | 650 | 3 | PCT-US94-1 Sequence 60, Applicati | 7.00e-21 |
| 16 | 362 | 7.3 | 724 | 1 | US-08-835- Sequence 62, Applicati | 1.88e-20 |
| 17 | 362 | 7.3 | 724 | 2 | US-08-121- Sequence 62, Applicati | 1.88e-20 |
| 18 | 362 | 7.3 | 724 | 3 | PCT-US94-1 Sequence 62, Applicati | 1.88e-20 |
| 19 | 362 | 7.3 | 724 | 1 | US-08-835- Sequence 62, Applicati | 1.88e-20 |
| 20 | 157 | 3.2 | 122 | 2 | US-08-835- Sequence 66, Applicati | 1.39e-03 |
| 21 | 157 | 3.2 | 122 | 1 | US-08-121- Sequence 66, Applicati | 1.39e-03 |
| 22 | 157 | 3.2 | 122 | 3 | PCT-US94-1 Sequence 66, Applicati | 1.39e-03 |
| 23 | 115 | 2.3 | 422 | 3 | PCT-US95-0 Sequence 170, Applicati | 1.67e+00 |

| | | | | | | |
|----|-----|-----|------|---|-----------------------------------|----------|
| 24 | 115 | 2.3 | 422 | 2 | US-08-469- Sequence 170, Applicat | 1.67e+00 |
| 25 | 115 | 2.3 | 422 | 2 | US-08-339- Sequence 3, Applicatio | 1.67e+00 |
| 26 | 115 | 2.3 | 422 | 3 | PCT-US94-0 Sequence 166, Applicat | 1.67e+00 |
| 27 | 115 | 2.3 | 422 | 2 | US-08-428- Sequence 3, Applicatio | 1.67e+00 |
| 28 | 115 | 2.3 | 422 | 2 | US-08-428- Sequence 3, Applicatio | 1.67e+00 |
| 29 | 115 | 2.3 | 422 | 2 | US-08-734- Sequence 170, Applicat | 1.67e+00 |
| 30 | 115 | 2.3 | 422 | 1 | US-08-036- Sequence 170, Applicat | 1.67e+00 |
| 31 | 115 | 2.3 | 422 | 1 | US-08-469- Sequence 170, Applicat | 1.67e+00 |
| 32 | 115 | 2.3 | 422 | 1 | US-08-249- Sequence 170, Applicat | 1.67e+00 |
| 33 | 115 | 2.3 | 422 | 1 | US-08-428- Sequence 3, Applicatio | 1.67e+00 |
| 34 | 115 | 2.3 | 422 | 1 | PCT-US94-0 Sequence 185, Applicat | 1.67e+00 |
| 35 | 115 | 2.3 | 422 | 2 | US-08-469- Sequence 170, Applicat | 1.67e+00 |
| 36 | 107 | 2.2 | 533 | 1 | US-08-445- Sequence 10, Applicati | 5.97e+00 |
| 37 | 107 | 2.2 | 533 | 1 | US-08-484- Sequence 13, Applicati | 5.97e+00 |
| 38 | 107 | 2.2 | 533 | 1 | US-08-484- Sequence 13, Applicati | 5.97e+00 |
| 39 | 100 | 2.0 | 780 | 1 | US-08-232- Sequence 14, Applicati | 1.77e+01 |
| 40 | 100 | 2.0 | 780 | 2 | US-08-786- Sequence 14, Applicati | 1.77e+01 |
| 41 | 100 | 2.0 | 1311 | 2 | US-08-340- Sequence 5, Applicatio | 1.77e+01 |
| 42 | 94 | 1.9 | 146 | 1 | US-07-795- Sequence 26, Applicati | 4.40e+01 |
| 43 | 94 | 1.9 | 146 | 1 | US-08-457- Sequence 26, Applicati | 4.40e+01 |
| 44 | 92 | 1.9 | 321 | 3 | PCT-US94-1 Sequence 26, Applicati | 5.93e+01 |
| 45 | 92 | 1.9 | 857 | 1 | US-07-717- Sequence 3, Applicatio | 5.93e+01 |

ALIGNMENTS

| RESULT | 1 | STANDARD: | PRT: | 771 AA. |
|-------------|--|------------|-------|-------------|
| ID | PCT-US94-10151A-54 | | | |
| XX | xxxxxx | | | |
| AC | | | | |
| DT | | | | |
| XX | | | | |
| DE | Sequence 54, Application PC/TUS9410151A | | | |
| XX | | | | |
| CC | GENERAL INFORMATION: | | | |
| CC | APPLICANT: The Regents of the University of California | | | |
| CC | TITLE OF INVENTION: The Semaphorin Gene Family | | | |
| CC | NUMBER OF SEQUENCES: 66 | | | |
| CC | CORRESPONDENCE ADDRESSES: | | | |
| CC | ADDRESS: FLEHR HOBBACH TEST ALBRITTON & HERBERT | | | |
| CC | STREET: 4 Embarcadero Center, Suite 3400 | | | |
| CC | CITY: San Francisco | | | |
| CC | STATE: CA | | | |
| CC | COUNTRY: USA | | | |
| CC | ZIP: 94111-4187 | | | |
| CC | COMPUTER READABLE FORM: | | | |
| CC | MEDIUM TYPE: Floppy disk | | | |
| CC | COMPUTER: IBM PC compatible | | | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | | |
| CC | SOFTWARE: Patentin Release #1.0, Version #1.25 | | | |
| CC | CURRENT APPLICATION DATA: | | | |
| CC | APPLICATION NUMBER: PCT/US94/10151A | | | |
| CC | FILING DATE: 13-SEP-1994 | | | |
| CC | CLASSIFICATION: | | | |
| CC | ATTORNEY/AGENT INFORMATION: | | | |
| CC | NAME: Osman, Richard A. | | | |
| CC | REGISTRATION NUMBER: 36,627 | | | |
| CC | REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO | | | |
| CC | TELECOMMUNICATION INFORMATION: | | | |
| CC | TELEPHONE: (415) 781-1989 | | | |
| CC | TELEFAX: (415) 398-3249 | | | |
| CC | TELEX: 910 277299 FHT UR | | | |
| CC | INFORMATION FOR SEQ ID NO: 54: | | | |
| CC | SEQUENCE CHARACTERISTICS: | | | |
| CC | LENGTH: 771 amino acids | | | |
| CC | TYPE: amino acids | | | |
| CC | TOPOLOGY: linear | | | |
| CC | MOLECULE TYPE: protein | | | |
| CC | SEQUENCE 771 AA: 88889 MW: 2976250 CN: | | | |
| Query Match | 13.3% | Score 661: | DB 3: | Length 771: |

| | |
|----|--|
| CC | GENERAL INFORMATION: |
| CC | APPLICANT: Goodman, Corey S. |
| CC | APPLICANT: Kolodkin, Alex L. |
| CC | APPLICANT: Matthes, David |
| CC | APPLICANT: Bentley, David R. |
| CC | APPLICANT: O'Connor, Timothy |
| CC | TITLE OF INVENTION: The Semaphorin Gene Family |
| CC | NUMBER OF SEQUENCES: 100 |
| CC | CORRESPONDENCE ADDRESS: |
| CC | ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP |
| CC | STREET: 268 Bush Street, Suite 3200 |
| CC | CITY: San Francisco |
| CC | STATE: CA |
| CC | COUNTRY: USA |
| CC | ZIP: 94104 |
| CC | COMPUTER READABLE FORM: |
| CC | MEDIUM TYPE: Floppy disk |
| CC | COMPUTER: IBM PC compatible |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS |
| CC | SOFTWARE: Patentin Release #1.0, Version #1.25 |
| CC | CURRENT APPLICATION DATA: |
| CC | APPLICATION NUMBER: US/08/121,713D |
| CC | FILING DATE: 13-SEP-1993 |
| CC | CLASSIFICATION: 514 |
| CC | ATTORNEY/AGENT INFORMATION: |
| CC | NAME: Osman, Richard A. |
| CC | REGISTRATION NUMBER: 36,627 |
| CC | REFERENCE/DOCKET NUMBER: B94-002-1 |
| CC | TELECOMMUNICATION INFORMATION: |
| CC | TELEPHONE: (415)343-4341 |
| CC | TELEFAX: (415) 343-4342 |
| CC | TELEX: |
| CC | INFORMATION FOR SEQ ID NO: 54: |
| CC | SEQUENCE CHARACTERISTICS: |
| CC | LENGTH: 771 amino acids |
| CC | TYPE: amino acid |
| CC | TOPOLOGY: linear |
| CC | MOLECULE TYPE: protein |
| CC | SEQUENCE 771 AA; 88889 MW; 2976250 CN; |
| CC | Query Match 13.3%; Score 661; DB 1; Length 771; |
| CC | Best Local Similarity 30.3%; Pred. No. 9.28e-47; Indels 47; Gaps 37; |
| CC | Matches 151; Conservative 126; Mismatches 174; |
| CC | Db 166 RGSYPDKLLTASLLIDELYSCTA-ADENGRGFALFRTLGHHNPIRTECHDSRWLNDP 224 |
| CC | QY 167 RGVAPFSDENSLVFEDEYVSTIRKCEYNGK-IPRRRRRGSELYTS-DIV-MQNP 222 |
| CC | Db 225 KFIASHLIESDNPRDVKVFFFRPMALIDGSHSKATHARIGQICKNDFGGHSL-VNKK 283 |
| CC | QY 223 QFIKATVHQ-DQADQIDKIYFFREDNDKKNPEARPLNYSRYAOLCRDGGESSLSYSKW 281 |
| CC | Db 284 TTFILKARLCSVPRNGIDTHFDELQVFLM-NFKDP-KNPVVYGVFTTSSNIEKSAVC 341 |
| CC | QY 282 NTFELKAMLVCS-DA-AINKNPNRIQDVFLLRDSGGMRDTRVYGVSPNPN-V--SAVC 335 |
| CC | Db 342 MYMSMDVRRVLEGPRAHNDGPNYQWVPRQGVHVPYPRGTCPSKTFGGFSDSTKDLRDDVIT 401 |
| CC | QY 336 VYSLGDDIDKVERITS-SLK-G--YH--S--S-LPNPRRGKCLRDQ-QPI-PRTEP-QV-A 381 |
| CC | Db 402 FARSHAPAMYNVEFPNNRPVIKTDVNYQFTQIVADVADBDGQ-YDVMETGIDVTVLK 460 |
| CC | QY 382 -DR-HPEVAQVVERPGRLKPL-FHSKYNHOKVAVNHRMQASHGETFHYLYLTDRGTINK 438 |
| CC | Db 461 VVSLPKFVWYDLEVLLEMTVFRPRTASLMELSTKQOOLYIGSTAGVQMLRPHRDDIY 520 |
| CC | QY 439 VVE-PGEDEHSAFENIME-IQFRRALAIQMSLDAERKLYVESQWSEVSOVPLDCEVY 496 |
| CC | Db 521 GKACAECLARDPYAMGASCSRYFPKARRKTRRQDRIKNDPLTHGSDLLHNDHGHSP 580 |
| CC | QY 497 GGGCHGCLMSRDPYCGMOCRCISYSE-KSYL-QSINAPRHPKPCPNRPDK-----AP 550 |
| CC | Db 581 EERIIVGVENSSTFLECSPKQARALVYWOFORNNEERKEETIRVDHIIIRIDQGLLSLQ 640 |

```

0Y      551 LQKVSIA-PNSRYVLYISC-PMESHATY-SW-FRKEVN-EQSCBPHQ-SPNCIFIEINLT 604
Db      641 QKDSGNYLCHAVEHGFQ 658
        :|::|||::|:
OY      605 AQOYGHYCEAOEGSYFR 622

RESULT   4
ID US-08-136-922-2 STANDARD: PRT: 477 AA.
AC xxxxxx
XX
XX
XX
DE Sequence 2, Application US/08136922
CC
CC Sequence 2, Application US/08136922
CC Patent No. 5416197
CC GENERAL INFORMATION:
CC APPLICANT: Kaper, Jonathan A.
CC APPLICANT: Luo, Yuling
CC TITLE OF INVENTION: Compositions Which Regulate Neural
CC TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
CC ADDRESSEE: No. 5416197ris
CC STREET: One Liberty Place
CC CITY: Philadelphia
CC STATE: PA
CC COUNTRY: USA
CC ZIP: 19103
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/136,922
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Deluca, Mark
CC REGISTRATION NUMBER: 33,229
CC REFERENCE/DOCKET NUMBER: UPN-1428
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-568-3100
CC TELEFAX: 215-568-3439
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 477 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
CC SEQUENCE 477 AA; 54647 MW; 1161988 CN;

Query Match 12.5%; Score 620; DB 1; Length 477;
Best Local Similarity 31.2%; Pred. No. 4,24e-43;
Matches 134; Conservative 110; Mismatches 145; Indels 41; Gaps 32;

Db      1 LEHDNEDDKYFFEFRENAIDGESHGAKATHAIGQICKNDFGGHRSI-VNKWTTFLLARL 59
OY      230 VHQDADAYDKIYYFFREDNDPKNEAPLANVSRAQLCRGDGOGSSLSYSKWNTFLXAML 289
Db      60 TCSVEGPANIDHFEDLOVFELM-NFKDP-KNVVYGVFTTSSNIIEFGSAVCMTSMSDVR 117
OY      290 VCS--DA-ATNNENFNLOVFLPLDPSGOWRDTRVYGVFSNPWN-Y--SACVYSLDID 343
Db      118 RYFLGPAHADBPNYVMWVYOGRVPRPQTCSKTGFGFDSTRKDLDPDVITTEFARSHPM 177
OY      344 KYFRIS-STK-G-YH-S-S-LPNRPFRCKCLPDQ-QPI-PIETTF--QV-A-DR-HPEV 387

```

Db 178 YNPVPMNNRPVITVDVNYGFOITVDVDAEDGQ-YDMVFIGDVGTVKVSIPKET 236
QY 388 AORVPEMPGRPLKTP-L-FHSKYNHYOKAAVYHRMOASHGEFPHVLYLTLDKGTIHKAYE-PGEQ 445
Db 237 WYDLEEVLLKEETVYFRETASISAMELSTKOOOLYIGSTAGYAOYLPHRCDIYKACAECC 286
QY 446 EHSFAFMIME-IQPRRAALOTSLDAERKKRLTVSSQWESVQVPLDCEYGGGCHGL 504
Db 297 LARDPYCAMDOSACSRYFPFLAKRRTRRDILRNGDPLTHCSDLHNDHNHSGPERRITGV 356
QY 505 MSRDYCGWDDGROIISYISSE-RSVL-QSINPAPRHKCEPNRPDK-----APLQKSLA 557
Db 357 ENSSTFLEECSPRSORALVYMQFORNREKREKEIVDHIITRTDGLLRSLOOKDSGNTL 416
QY 558 PMSRYVLYSC-EMERHATY-SW-RHKENV-EQSCPEGHQ-SPNCIFILENTLTAQOYGHYE 612
Db 417 CHAYEHGFIQ 426
QY 613 CEAOEGSYER 622

```
Db      115 LVCGTNGNPGCMK-IDSSDPKHGRGYARAYQMSKYTIISHNGC-VLSDINISK-EG-I 170
        I | | | | : | | : | | : | | | | | | | | | | | | | | | | | | | | | | |
RESULT ID PCT US94-10151A-56 STANDARD; PRF; 441 AA.
XX AC xxxxxx
XX DT
XX DE Sequence 56, Application PC/TUS9410151A
CC CC
CC Sequence 56, Application PC/TUS9410151A
CC GENERAL INFORMATION:
CC APPLICANT: The Regents of the University of California
CC TITLE OF INVENTION: The Semaphorin Gene Family
CC NUMBER OF SEQUENCES: 66
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
CC STREET: 4 Embarcadero Center, Suite 3400
CC City: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111-4187
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/10151A
CC FILING DATE: 13-SEP-1994
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Osman, Richard A.
CC REGISTRATION NUMBER: 36,627
CC REFERENCE/DOCKET NUMBER: FP-56750-PC/RAO
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC TELEX: 910 277299 FHT UR
CC INFORMATION FOR SEQ ID NO: 56:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 441 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
CC CC
CC SEQUENCE 441 AA; 50185 MW; 1082834 CN;
```

| | | | |
|----|-----|---|-----|
| Qy | 141 | LACGNNARHPSQWMLVNGTVVP/LGEMRGYADPSPDENSILVFEGDEYVSTRKQENYGI | 200 |
| Db | 171 | KWRREFDPCGYDLTYADNAVIPKDG-LRGAFVDRKQGYD-KVYILFTDTIGSR--I-VK | 225 |
| Qy | 201 | PFRRIRGE-S-ELTSTPTWQNPQFIKATIVHDQAVDKRIYFFEDNPDKNPEAPLN | 258 |
| Db | 226 | IPYIKOMLGNDEGSSLSHRSMTFLKVELC-DID-GRSY-R-Q-IHSRTIKTD-ND | 279 |
| Qy | 259 | VSRAVQALRGDQGGESSLSYSKWMTFLKAMLYCSDAATNNFNRLQVFLPDPGQWRD | 318 |
| Db | 280 | TILYFEDSPYSKALCTYSNMITKQSFSTKLEGYKOLPSPASGICLPAGVAVHTPF | 339 |
| Qy | 319 | TRYGVGEFNPWNYSAVCYSLSDGIDKFRIRISLKGTHSSLPNRPQGCAPDQPIPTFE | 378 |
| Db | 340 | EVIEKYANVLDLITPLS | 356 |
| Qy | 379 | QVADRHPEVAQRAPEPMG | 395 |

| RESULT | 6 | STANDARD; | PRT; | 441 AA. |
|--------|--|-----------|------|---------|
| ID | US-08-835-268-56 | | | |
| XX | xxxxxx | | | |
| AC | | | | |
| XX | | | | |
| DT | | | | |
| XX | | | | |
| DE | Sequence 56, Application US/08835268 | | | |
| XX | | | | |
| CC | Sequence 56, Application US/08835268 | | | |
| CC | Patent No. 5807826 | | | |
| CC | GENERAL INFORMATION: | | | |
| CC | APPLICANT: Goodman, Corey S. | | | |
| CC | APPLICANT: Kolodkin, Alex L. | | | |
| CC | APPLICANT: Mattes, David L. | | | |
| CC | APPLICANT: Bentley, David R. | | | |
| CC | APPLICANT: O'Connor, Timothy | | | |
| CC | TITLE OF INVENTION: The Semaphorin Gene Family | | | |
| CC | NUMBER OF SEQUENCES: 100 | | | |
| CC | CORRESPONDENCE ADDRESS: | | | |
| CC | ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP | | | |
| CC | STREET: 268 Bush Street, Suite 3200 | | | |
| CC | CITY: San Francisco | | | |
| CC | STATE: CA | | | |
| CC | COUNTRY: USA | | | |
| CC | ZIP: 94104 | | | |
| CC | COMPUTER READABLE FORM: | | | |
| CC | MEDIUM TYPE: Floppy disk | | | |
| CC | COMPUTER: IBM PC compatible | | | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | | |
| CC | SOFTWARE: PatentIn Release #1.0, Version #1.25 | | | |
| CC | CURRENT APPLICATION DATA: | | | |
| CC | APPLICATION NUMBER: US/08/835,268 | | | |
| CC | FILING DATE: | | | |
| CC | CLASSIFICATION: | | | |
| CC | PRIOR APPLICATION DATA: | | | |
| CC | APPLICATION NUMBER: US/08/121,713 | | | |
| CC | FILING DATE: 13-SEP-1993 | | | |
| CC | ATTORNEY/AGENT INFORMATION: | | | |
| CC | NAME: Osman, Richard A. | | | |
| CC | REGISTRATION NUMBER: 36,627 | | | |
| CC | REFERENCE/DOCKET NUMBER: B94-002-1 | | | |
| CC | TELECOMMUNICATION INFORMATION: | | | |
| CC | TELEPHONE: (415) 343-4341 | | | |
| CC | TELEFAX: (415) 343-4342 | | | |
| CC | TELEX: | | | |
| CC | INFORMATION FOR SEQ ID NO: 56: | | | |
| CC | SEQUENCE CHARACTERISTICS: | | | |
| CC | LENGTH: 441 amino acids | | | |
| CC | TYPE: amino acid | | | |
| CC | TOPOLOGY: linear | | | |
| CC | MOLECULE TYPE: protein | | | |
| CC | SEQUENCE 441 AA: 50185 MW; 1082834 CN; | | | |
| CC | 50 | | | |

| | |
|----|--|
| DE | Sequence 56, Application US/08121713D |
| xx | |
| CC | Sequence 56, Application US/08121713D |
| CC | Patent No. 5639856 |
| CC | GENERAL INFORMATION: |
| CC | APPLICANT: Goodman, Corey S. |
| CC | APPLICANT: Kolodkin, Alex L. |
| CC | APPLICANT: Mathes, David |
| CC | APPLICANT: Bentley, David R. |
| CC | APPLICANT: O'Connor, Timothy |
| CC | TITLE OF INVENTION: The Semaphorin Gene Family |
| CC | NUMBER OF SEQUENCES: 100 |
| CC | CORRESPONDENCE ADDRESS: |
| CC | ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP |
| CC | STREET: 268 Bush Street, Suite 3200 |
| CC | CITY: San Francisco |
| CC | STATE: CA |
| CC | COUNTRY: USA |
| CC | ZIP: 94104 |
| CC | COMPUTER READABLE FORM: |
| CC | MEDIUM TYPE: Floppy disk |
| CC | COMPUTER: IBM PC compatible |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS |
| CC | SOFTWARE: PatentIn Release #1.0, Version #1.25 |
| CC | CURRENT APPLICATION DATA: |
| CC | APPLICATION NUMBER: US/08/121,713D |
| CC | FILING DATE: 13-SEP-1993 |
| CC | CLASSIFICATION: 514 |
| CC | ATTORNEY/AGENT INFORMATION: |
| CC | NAME: Osman, Richard A. |
| CC | REGISTRATION NUMBER: 36,627 |
| CC | REFERENCE/DOCKET NUMBER: B94-002-1 |
| CC | TELECOMMUNICATION INFORMATION: |
| CC | TELEPHONE: (415)343-4341 |
| CC | TELEFAX: (415) 343-4342 |
| CC | TELEX: |
| CC | INFORMATION FOR SEQ ID NO: 56: |
| CC | SEQUENCE CHARACTERISTICS: |
| CC | LENGTH: 441 amino acids |
| CC | TYPE: amino acid |
| CC | TOPOLOGY: linear |

DE Sequence 58, Application US/08835268
XX
CC Sequence 58, Application US/08835268
CC Patent No. 5807826
CC GENERAL INFORMATION:
CC APPLICANT: Goodman, Corey S.
CC APPLICANT: Kolodkin, Alex L.
CC APPLICANT: Mathies, David
CC APPLICANT: Bentley, David R.
CC APPLICANT: O'Connor, Timothy
CC TITLE OF INVENTION: The Semaphorin Gene Family
CC NUMBER OF SEQUENCES: 100
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
CC STREET: 268 Bush Street, Suite 3200
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/835,268
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/121,713
CC FILING DATE: 13-SEP-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Osman, Richard A.
CC REGISTRATION NUMBER: 36,627
CC REFERENCE/DOCKET NUMBER: B94-002-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415)343-4341
CC TELEFAX: (415) 343-4342


```

CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US94/10151A
CC      FILING DATE: 13-SEP-1994
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Osman, Richard A.
CC      REGISTRATION NUMBER: 36,627
CC      REFERENCE/DOCKET NUMBER: EP-58750-PC/RAO
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (415) 781-1989
CC      TELEFAX: (415) 398-3249
CC      TELEX: 910 277299 FHT UR
CC      INFORMATION FOR SEO ID NO: 64:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 712 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 712 AA; 79751 MW; 2715572 CN;

Db      Query Match 7.6%; Score 378; DB 3; Length 712;
      Best Local Similarity 27.0%; Pred. No. 7,91e-22;
      Matches 140; Conservative 128; Mismatches 194; Indels 56; Gaps 43.

Db      48 HFTVLYNO-DEMSILVGGNRRYNIIPDLSRKGGRIDWPS-SDAHQGLCLTKKTDDC 105
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      75 HTVLEHGGSSVWVGGRKYL-L-IDPEGKNAVSRTYVIGSTKS-C-L-DKRP-C 126
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      106 QNYRIILYSSSEPGKLVIGTINSYPLCTRYAFKKGKYLVEKEVEGIGLCEPYNEHNSTV 165
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      127 ENYTLILERSBG-LLAGCTNARHPSGCMNLV-NGT-VPLG-EMRGYAPSPDENSL-V 180
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      166 SYNG-QLEFSATV-ADESGCDPLIYREPORTL-SDLKOLNAPNFVNS--V---AYGDI 216
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      181 LEEGDEVYSTIRKOEYNKIPREFRIRISESLYSTDYVMPQPIKATIVHQDAYDKI 240
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      217 FEFYRETAVEYMNCGKYL-Y-SRVARVCKDDKG-PSHOSRDRWTSFLKARLNCSTPGEPF 274
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      241 YFFREDNPD-KNEPAPLANSRVKQLCRGDGGGSSLSYSKMTFLKAMLVCSDAATNK- 298
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      275 YFDEIQSTSDIVEGRYNSDDSKITLYGILTPVAIGSALCAVQAMADILVEGSEFKHQ 334
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      299 NFNRLQDVFLLPDSGQQRDIR--VYGFNSPMN-Y--SACVYSLDDIDKVF----R-T 348
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      335 ETINSNMLVPONTLVPEPRPGOCVRSRILPDKNVNFIKTHSLMED-VPALFG-KPVLVR 392
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      349 SSLGK-Y-H-S-SL-PNRPFGKCLPDQPIPTFTFFQVADRHPVAAQREVEPMGLKPLFH 403
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      393 VSLQRFALIVDPQVKTINNQYLDLYLIGDDCKVYLKAVNIPIRKHAKALLYKRYFSVH 452
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      404 -S-KTHYQKAVH-RMQASHGETHVIYLLTDRGTIKHYVE-PDEQHSFAFINMETQPF 459
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      453 PHGAPVKQLKAPGKGVVVGKDEIRLANPHCAS-KTRCKDCVDELQDPHCAMDAKONL 511
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      460 RRAALIQMSLDAERKRIYSSQMEVSVPYLDLCEYVGGGCHGCLMSRDPYCGMD-QGR 517
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      512 CVSIDYVTSYRFLIQDVVGRDGNK-CWSPQTDKRTVIK 548
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      518 CISIYS-SE-RSVLQSIINPAPRHKECPNPKRDKAPLQK 553
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
ID US-08-121-713D-60 STANDARD; PRT; 650 AA.
AC xxxxxx
XX
XT
XX
Sequence 60, Application US/08121713D

```

| | | |
|----|-----|--|
| Cc | XX | Sequence 60, Application US/08121713D |
| Cc | | Patent No. 5639856 |
| Cc | | GENERAL INFORMATION: |
| Cc | | APPLICANT: Goodman, Corey S. |
| Cc | | APPLICANT: Kojodkin, Alex L. |
| Cc | | APPLICANT: Matthes, David |
| Cc | | APPLICANT: Bentley, David R. |
| Cc | | APPLICANT: O'Connor, Timothy |
| Cc | | TITLE OF INVENTION: The Semaphorin Gene Family |
| Cc | | NUMBER OF SEQUENCES: 100 |
| Cc | | CORRESPONDENCE ADDRESS: |
| Cc | | ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP |
| Cc | | STREET: 268 Bush Street, Suite 3200 |
| Cc | | CITY: San Francisco |
| Cc | | STATE: CA |
| Cc | | COUNTRY: USA |
| Cc | | Zip: 94104 |
| Cc | | COMPUTER READABLE FORM: |
| Cc | | MEDIUM TYPE: Floppy disk |
| Cc | | COMPUTER: IBM PC compatible |
| Cc | | OPERATING SYSTEM: PC-DOS/MS-DOS |
| Cc | | SOFTWARE: PatentIn Release #1.0, Version #1.25 |
| Cc | | CURRENT APPLICATION DATA: |
| Cc | | APPLICATION NUMBER: US/08/121,713D |
| Cc | | FILING DATE: 13-SEP-1993 |
| Cc | | CLASSIFICATION: 514 |
| Cc | | ATTORNEY/AGENT INFORMATION: |
| Cc | | NAME: Osman, Richard A. |
| Cc | | REGISTRATION NUMBER: 36,627 |
| Cc | | REFERENCE/DOCKET NUMBER: B94-002-1 |
| Cc | | TELECOMMUNICATION INFORMATION: |
| Cc | | TELEPHONE: (415)343-4341 |
| Cc | | TELEFAX: (415) 343-4342 |
| Cc | | TELEX: |
| Cc | | INFORMATION FOR SEQ. ID NO: 60: |
| Cc | | SEQUENCE CHARACTERISTICS: |
| Cc | | LENGTH: 650 amino acids |
| Cc | | TYPE: amino acid |
| Cc | | TOPOLOGY: linear |
| Cc | | MOLECULE TYPE: protein |
| Cc | | SEQUENCE 650 AA; 72940 MW; 2193067 CN; |
| Cc | | |
| Cc | | Query Match 7.4%; Score 367; DB 1; Length 650; |
| Cc | | Best Local Similarity 28.4%; Pred. No. 7.00e-21; |
| Cc | | Matches 124; Conservative 104; Mismatches 158; Indels 51; Gaps 38; |
| Dd | | |
| Dd | 3 | DCQNTIRMTVVPSPGRLEFCVGCTNSFRPMCNTYIISDSNLYLTKNGCAQCPIDPRHNST 62 |
| Dd | 125 | DCENVTILLERRSG-LPLACGTNARHPSQWN-LVNGTVPLGEMR-GYA--PSPDENSL 179 |
| Dd | 63 | SVLADNELYSTV-ADSSGSDDPIYREPLDTQEYDIS-LNAPMFV-SSPT-OGD-F--- 114 |
| Dd | 180 | VLEDEDEYYSITRKQENKGIIPRRRIARGSELYTSDTVMNOPQFIATIVHDQAYDDK 229 |
| Dd | 115 | VYFFRETAVAEFINCGKAIFY-SRYARCYKMWDGGPHFR-NRMTSFLKSLNCISPGDY 172 |
| Dd | 240 | IYYEFRENPD-KNPPEARLVNSRAOICLRDQGEGESSLSVKNNTFILKALVCSDAITNK 238 |
| Dd | 173 | FYNELIGASNLVEGYGSMSSKLIIYVENMTPSNSIPGSAVCAPAIODIADTEFGOFKEO 232 |
| Dd | 299 | NFNRLDQVF-LLPDPGSGWMHRTGVYGFNSPMW-Y--SAVCYYSLODI-DKV---FR-- 347 |
| Dd | 223 | TGISNMPLPVNNNAVPPRPBGSCINDSRALPDLPLNKTIKHSIMDENVAFPFSQPIIVRT 232 |
| Dd | 348 | TS--S--LKGYHSSLPNRPBPKCLPDQOPIPTETFEYVADRHNPVAAQVPEMPGLKTPLFH 403 |
| Dd | 293 | STIREFOIAVDADIKTPGKTYDIVIGTDHGKIISKVAESADSADKVTSVIEEDIV 352 |
| Dd | 404 | SK-IHYQKAVNH-RMQASHGETHFVILTLITDRGTHK-V-VEGE-QEH--S--FA-FNI 453 |
| Dd | 353 | LTKSEPIRNLETIVTMQYDQPKDGSYDDGKLIIVTDSQVVAIOLHRCHNDKITSCSECV 412 |

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mprch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 3 16:24:21 1999; Maspar time 24.64 Seconds

Tabular output not generated. 574,766 Million cell updates/sec

Title: >US-09-240-410-2

Description: (1-666) from US09240410.pep

Perfect Score: 4968

Sequence: 1 MTRPPGGRAPSAAPRARVPG.....LAASIMLGVLPTLTGLLVH 666

Scoring table: PAM 150

Gap 11

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-genesq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 37.430; Variance 164.579; scale 0.227

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query Length | DB ID | Description | Pred. No. |
|------------|-------|-------|--------------|-------|------------------------------|-----------|
| 1 | 661 | 13.3 | 771 | 13 | R1380 Human semaphorin III | 1.46e-48 |
| 2 | 620 | 12.5 | 477 | 13 | R7415 Human collapsin. | 1.01e-44 |
| 3 | 622 | 12.5 | 775 | 33 | W63748 Human semaphorin. | 6.59e-45 |
| 4 | 606 | 12.2 | 861 | 22 | W58540 Mouse CD100 antigen. | 2.09e-43 |
| 5 | 599 | 12.1 | 861 | 32 | W58540 Human semaphorin. | 9.30e-43 |
| 6 | 577 | 11.6 | 776 | 32 | W51313 Rat semaphorin W. | 1.05e-43 |
| 7 | 544 | 11.0 | 862 | 22 | W17657 Human CD100 antigen. | 1.23e-47 |
| 8 | 463 | 9.3 | 441 | 13 | R1381 Vaccinia virus semaph. | 3.70e-30 |
| 9 | 459 | 9.2 | 730 | 13 | R1379 Grasshopper semaphori | 8.62e-30 |
| 10 | 428 | 8.6 | 587 | 32 | W51314 Human semaphorin W. | 5.91e-27 |
| 11 | 390 | 7.9 | 974 | 33 | W64221 Human secreted protei | 1.68e-23 |
| 12 | 378 | 7.6 | 712 | 13 | R1384 Tribolium semaphorin | 2.04e-22 |
| 13 | 367 | 7.4 | 650 | 13 | R1382 Drosophila semaphorin | 2.00e-21 |
| 14 | 362 | 7.3 | 724 | 13 | R1383 Drosophila semaphorin | 5.63e-21 |
| 15 | 352 | 7.1 | 930 | 32 | W57260 Human semaphorin Y. | 4.45e-20 |
| 16 | 325 | 6.5 | 888 | 25 | W19857 Human semaphorin Z. | 1.14e-17 |

| | | | | | | |
|----|-----|-----|------|----|------------------------------------|----------|
| 17 | 316 | 6.4 | 887 | 25 | W19856 Rat semaphorin Z. | 7.19e-17 |
| 18 | 296 | 6.0 | 929 | 32 | W57259 Human semaphorin Y. | 4.19e-15 |
| 19 | 157 | 3.2 | 122 | 13 | R1385 Varicella major virus s | 2.14e-03 |
| 20 | 114 | 2.3 | 111 | 32 | W51315 Human semaphorin W pa | 3.76e+00 |
| 21 | 114 | 2.3 | 248 | 17 | R87445 Human glial growth fa | 3.19e+00 |
| 22 | 115 | 2.3 | 248 | 21 | W09358 Human glial growth fa | 3.19e+00 |
| 23 | 115 | 2.3 | 248 | 17 | R96074 Human glial growth fa | 3.19e+00 |
| 24 | 115 | 2.3 | 248 | 9 | R46912 GGF segment E. | 3.19e+00 |
| 25 | 115 | 2.3 | 248 | 9 | R55653 GGF segment E. | 3.19e+00 |
| 26 | 112 | 2.3 | 248 | 13 | R67242 Human glial cell gro | 5.22e+00 |
| 27 | 116 | 2.3 | 422 | 17 | R87467 Human glial growth factor s | 2.70e+00 |
| 28 | 115 | 2.3 | 422 | 21 | W09372 Human GGF2. | 3.19e+00 |
| 29 | 115 | 2.3 | 422 | 17 | R96081 Glial growth factor G | 3.19e+00 |
| 30 | 115 | 2.3 | 422 | 13 | R67258 Human glial cell grow | 3.19e+00 |
| 31 | 115 | 2.3 | 422 | 17 | R86678 Mature hGF2. | 3.19e+00 |
| 32 | 115 | 2.3 | 422 | 21 | W09371 Human neuropilin GGF2 | 3.19e+00 |
| 33 | 115 | 2.3 | 422 | 9 | R46923 GGF-II encoded by clo | 3.19e+00 |
| 34 | 115 | 2.3 | 422 | 9 | R55654 GGF-II encoded by clo | 3.19e+00 |
| 35 | 115 | 2.3 | 422 | 17 | R87466 Glial growth factor s | 3.19e+00 |
| 36 | 98 | 2.0 | 176 | 24 | W14557 Streptococcus pneumon | 4.87e+01 |
| 37 | 99 | 2.0 | 591 | 4 | R23006 Protein transcribed f | 4.17e+01 |
| 38 | 100 | 2.0 | 780 | 28 | W47039 Soluble truncated VEG | 3.57e+01 |
| 39 | 100 | 2.0 | 780 | 13 | R62487 Truncated FLT SVGF-R | 3.57e+01 |
| 40 | 98 | 2.0 | 832 | 33 | W61092 Tag DNA polymerase I | 4.87e+01 |
| 41 | 98 | 2.0 | 1091 | 27 | W41641 Sequence used in dete | 4.17e+01 |
| 42 | 98 | 2.0 | 1093 | 9 | R56979 Human myotonic dystro | 4.87e+01 |
| 43 | 101 | 2.0 | 2329 | 25 | W25080 Partial BRCA2 cancer | 3.05e+01 |
| 44 | 96 | 1.9 | 832 | 33 | W61090 Tag DNA polymerase I | 6.64e+01 |
| 45 | 95 | 1.9 | 2965 | 31 | W56450 Fragment F1029 of a n | 7.74e+01 |

ALIGNMENTS

RESULT 1
 ID R1380 standard; Protein; 771 AA.
 AC R71380;
 DT 21-NOV-1995 (first entry).
 DE Human semaphorin III protein.
 KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
 KW varicella major virus; smallpox; semaphorin receptor binding activity;
 KW modulation; nerve cell growth; immune response; viral pathogenesis;
 KW neurodegenerative disease; neuro-regeneration; oncological infection.
 OS Homo sapiens.
 PN W09507706-A.
 PD 23-MAR-1995.
 PE 13-SEP-1994; U10151.
 PR 13-SEP-1993; US-121713.
 PA (REGC) UNIV CALIFORNIA.
 PI Bentley DR, Goodman CS, Kolodkin AL, Mathes D;
 PI O'Connor T;
 DR WPI; 95-131177/17.
 DR N-PSDB: Q87442.
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2: Page 60-63; 101pp; English.
 CC The sequence of the human semaphorin III protein.
 CC encoded by the grasshopper semaphorin III (Q87441), human semaphorin III,
 CC vaccinia virus semaphorin IV (Q87443), Drosophila semaphorin I and II,
 CC (Q87444-5), Tribolium semaphorin I (Q87446) or varicella major (smallpox)
 CC virus semaphorin IV (Q87447) genes were used to generate a series of
 CC peptides (R70370-R70418), which retain semaphorin receptor binding
 CC activity. The semaphorin derived or semaphorin receptor derived peptides
 CC are potent modulators of nerve cell growth, immune responsiveness and
 CC viral pathogenesis. They can be used in diagnosis and treatment of
 CC neurological disease and neuro-regeneration, immune modulation and
 CC diagnosis and treatment of viral and oncological infection and diseases.
 SO Sequence 771 AA;
 Query Match 13.3%; Score 661; DB 13; Length 771;
 Best Local Similarity 30.3%; Pred. No. 1.46e-48;
 Matches 151; Conservative 126; Mismatches 174; Indels 47; Gaps 37;
 Db 166 rgsxpydpklltallldgslgysgta-adfmgdftaifrlghpirtteghsrwvndp 224

| | | | |
|----|-----|--|------|
| Oy | 167 | ROYAFSPDENLUVFEGDEVIYSTRKOEYINK-IPFRRIIRGESELYTS--DTV-WQNP | 2222 |
| Oy | 225 | KfIsaHIsesInpedkdVkyvffIenaIdgeHshkatharIgdIckndfgHzal-vnK | 2833 |
| Oy | 223 | QfIKATIvHQ-DOAVDDKIYFFEFREDNDPKNDEAPLNVSRVAOLCRDQOGESLSYSKM | 281 |
| Db | 284 | tfllkarIcsvprpnqIdtHfdelqdvflm-nfKdp-knpvuygvfItssnfKgsavc | 3411 |
| Oy | 282 | ntfELKAMLVCS--DA-ATKNKNENRQDVELLPDPSGGWRDRYGVGFSPNPN-Y--SAYC | 3353 |
| Db | 342 | mysmsdvrrvflgryahhdgpnnyqvvrygvprprrptcspkffgfsdskdipddvIt | 4010 |
| Oy | 336 | VYSLSDIDKVRFTS-SLK-G-YH-S-S-LPmPRpEckLPDO-QPI-PlETfF-QV-A | 3811 |
| Db | 402 | fashpamyhpyfpmnnpIvIktdvnyqfclqIvvdvdaedqg-yvmtffIgdvrvIvK | 460 |
| Oy | 382 | -DR-HPEVAQRvERpGRpKLTPL-FHSKHNyOKAVHNHQAshGTfHVLVLTTRDGRtHK | 438 |
| Db | 461 | vvsIpKetwYdLeevlIeemtvfIeptaIsameIsckqgqlygstagvaqlpHbcdIy | 520 |
| Oy | 439 | VVE-PGEDEHSFAFNIME-IOPFRRAALIQTMSLDARERKLYVSSQMEVSQVPLDCEVY | 496 |
| Db | 521 | qKacaeccIasdpYcawgsacsYufbKarkrttrrgdIrngdrlthscdlhnanhghsp | 560 |
| Oy | 497 | GGGCGGCLMSMDPFCGMQDQRCISITySE-NSVY-QSINPvEPREKCPNKPpK---AP | 550 |
| Db | 581 | eerIlyyvensstfIeCspxsqarIyvWqfgrnneekelIrvddhIiirtdggllIrsId | 640 |
| Oy | 551 | LQKVSIA-PNSRYTLSC-PMSRRATY-SW-RHKNvN-EDSCRGHIO-SNCLITfIENLT | 604 |
| Db | 641 | qKdsGnylchavehgfId 658 | |
| Oy | 605 | AQOYGHYfCEAOEGSYR 622 | |

[illegible]

| QY | 230 | VHDDQAYDKDKIYFFFEEDNPDKNPEAPLVNSRYAQLCRDDGCGESSLSYKNNTEFLKAML | 289 |
|-----------------------|--|--|-----------------|
| Db | 60 | icsvpngidcthdclgcvflm-nfkdp-knpvvygvttsnlfkgsavcmysmdvr | 117 |
| QY | 290 | VCS--DA-ATNNKFNRLADVFLLPDPSGQWBDTRVYGVSFNPWN-Y--SACVYSLGIDID | 343 |
| Db | 118 | ryfllgpyahrdpnygwvyggrvpyprgtpctpsktffgfsdtkalpdvlfarfshpam | 177 |
| QY | 344 | KVFRTS-SLK-G-YH-S-S-LPMPRGKCLPQO-OP1-PTETP--OV-A-DR-HPEV | 387 |
| Db | 178 | ynpvmfmmrprlyfktcdvnygftglvdvdadeedg-ydvymcflgdvgyvlkvpslpket | 236 |
| QY | 388 | AQRVEPMGSLKPLP-FHSXTHQKVAVHRMQASHSTEFHVLTLTTRDGTIRKHYVE-PEDQ | 445 |
| Db | 237 | wydlleevllaeentvtfreptalsamelstckqgllygstagaqpldhrcdiylgkacaecc | 296 |
| QY | 446 | EHSEFANIME-IQPRRRAAIOTMSLDNERRLLYSSQMEVSOVPLDCEVYGGGCHGL | 504 |
| Db | 297 | lardyfcawdgsaacsryfptakrttrirgdlngdplthcsdlhnhhnghspeerilygv | 356 |
| QY | 505 | MSRDPYCGMDQRCISYSSSE-RSVL-QSINPAEPHKCECPNPKPDK---APLQKVSILA | 557 |
| Db | 357 | enstffleospksqralywygqgrneekkeirvddhntirdgglllrsjqksgsnyl | 416 |
| QY | 558 | PMSRYTLC-PESHRTATY-SW-RHKENV-EGSCPEHQ-SPNCILFENETLTAQYGHXF | 612 |
| Db | 417 | chaveghfag 426 | |
| QY | 613 | CEAOGGSYFR 622 | |
| RESULT | 3 | | |
| ID | W63748 | standard; Protein; 775 AA. | |
| AC | W63748; | | |
| DI | 01-OCT-1998 | (first entry) | |
| DE | Human semaphorin. | | |
| KW | Human; semaphorin; diagnosis; nerve disease; nerve growth inhibitor; | | |
| KW | neurological disease; atopic skin inflammation; autoimmune disease; | | |
| KW | pain. | | |
| OS | Homo sapiens. | | |
| PN | MO9822504-AL. | | |
| PD | 28-MAY-1998. | | |
| PR | 12-NOV-1997; J04111. | | |
| PR | 12-NOV-1996; JP-321068. | | |
| PA | (SUMU) SUMITOMO PHARM CO LTD. | | |
| PI | Furuyama T, Inagaki S; | | |
| DR | WPI: 98-312416/27. | | |
| DR | N-PSDB: V35367. | | |
| PT | Gene encoding new semaphorin nerve growth inhibitor - useful in | | |
| PT | diagnosis, treatment and study of neurological diseases | | |
| PS | Claim 1; Page 33-37; 49pp: Japanese. | | |
| CC | The present sequence represents human semaphorin, a nerve growth | | |
| CC | inhibitor. The semaphorin protein, and gene encoding the protein, | | |
| CC | and their derivatives, are used in the diagnosis, treatment and | | |
| CC | study of neurological disorders such as atopic skin inflammation, | | |
| CC | autoimmune diseases and pain. | | |
| SQ | Sequence 775 AA; | | |
| Query Match | 12.5%; | Score 622; | DB 33; |
| Best Local Similarity | 30.3%; | Pred. No. 6.59e-45; | |
| Matches | 169; | Conservative 139; | Mismatches 188; |
| | | Indels 62; | Gaps 45; |
| Db | 2 | apaghi-ftllwphllelwyprgsanpysprllstskellelnrtstfgepjlfg--ldl 58 | |
| QY | 20 | GPRNGLGPLRLR-LTLLLMWMAASAO-G-HLKSGR-IFAWWKGVGDRODRDFGOTEP 74 | |
| Db | 59 | htmlldcy-qerlitvggrdlyvsnlegrvdegyrelywptstavyveecmkjgkxanecan 117 | |
| QY | 75 | HTVLFHEGGSSVWVGGRGKYV-L-PD-FEFG-KNA--SVRTVWIGSTKSSCLDKRCCEN 128 | |
| Db | 118 | ytyvlhyhntthlltcatgaefdrhcafiyvgbhseepdlfheshreerggrgrefdpns 177 | |
| QY | 129 | YITLIER-RSEGLLACTGNARHPSC-WNLV-NGTVVPL-----GEM-RGYAPFSPDEN 177 | |


```

QY      125 DCENVITLLER-RSEGLIACGTNARHPSG--WNLVNGTVVPLGEM-KGIYAFSPDSENSLV 180
Db      167 vmvvgelysgts-ynfigsepilrnsnshpdlrteyalpwnpsfvfadviqspdpge 225
QY      181 LFEDEVYSTIRKOEYNGKIPFRFRIRIGESSELYTSDTV--MÖNQFIKATIVH-Q-D-Q 234
Db      226 geddkvyfftevevefvefklmipvarvckgdggltlq-klwtstlkarllscxp 284
QY      235 AYDKIYFFREDNPÖNPEAPLVNSVAQICRGDGGESSLSVSKMNTFLKAMLVCSDA 294
Db      285 dsglvfnildvfvlar-p-gl-kepvfyavftpqlnnvgsavcaaylacveavfsrgky 342
QY      295 ATKNNFNRLQDVFLPDPSPGQMDTRRYGVFSNPNMNT---SAYCVYSLGIDKVF-R-- 347
Db      343 mgsatveqgsthtkwyryngpvrtpgacidsaaraanytslnlpkltlqfvkdhplmd 402
QY      348 --TSSL---K-G-YHSSLPNRPCKLT-PDQQPIR-TETFOVADRNPE-V-AQR-VE- 392
Db      403 svrpidmrplkikdvnyqivdvrtgaldgtfydmfistdgalhkvavlkveh-v- 460
QY      393 PMGPKLT-P-LFHSKHYHOKVAVNHRQASHGETFHVLYLTDRGTHKVVPEGEQHSFA 450
Db      461 --ieetqlfrdfevrltlllskkyrkfyasngsvvgaπλαcekhgs-ceedvcland 517
QY      451 FNMIEIOPFRRAAIQTMSTDAER-RKL-YVSQWEVSQVPLDLCEVYGGCGHGLMSRD 508
Db      518 PYCawspalkacvclthgeas 538
QY      509 PYCGMDQ-G-R-CISTYSSERS 527

```

RESULT 6
ID W51313 standard; Protein: 776 AA.

```

AC W51313;
DT 08-SEP-1998 (first entry)
DE Rat semaphorin W.
KW Rat; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
KW Immunosuppressant; gene therapy; diagnosis; research reagent.
PN M09815628-A1.
PD 16-APR-1998.
PF 03-OCT-1997; J03549.
PR 09-OCT-1996; JP-287636.
PA (SUMU ) SUMITOMO PHARM CO LTD.
PI Kikuchi K, Kimura T;
DR WPI: 98-261015/23.
DR N-PSDB: V07279, V07280.
PT Nerve extension inhibitor protein semaphorin W - is useful as
PT therapeutic drug and diagnostic and research reagent
PS Claim 1; Page 60-64; 90pp; Japanese.
CC The present sequence represents rat semaphorin W. Semaphorin W and
CC its derivatives are nerve extension inhibitors which are useful as
CC antiallergic, immunosuppressant and anticancer agents. The DNA
CC encoding semaphorin W can also be used in gene therapy, e.g. using
CC a viral vector. The proteins, peptides, DNA and antibodies which
CC recognise the protein or peptides, can be used as diagnostic or
CC research reagents. Semaphorin W can be used as a screen for
CC semaphorin W antagonists with possible therapeutic use.
SQ Sequence 776 AA;

```

Query Match 11.6%; Score 577; DB 32; Length 776;
Best Local Similarity 27.8%; Pred. NO. 1,05e-40;

Matches 171; Conservative 149; Mismatches 240; Indels 56; Gaps 40;

```

Db      5 aerprgprppvfpfpfppllllllaasavcgrvpsvprtsplseadylrrfca 64
QY      13 APRARV-PCGPARTGLPLRLRLLLLMAAASQGH-HSGPRI-FAWKGHVGDQDVDF 69
Db      65 shlynyallvdpashtlylgardsifaltlpsgerprtidmvp-ethrgncrckgk 123
QY      70 GOTEPTVLEHPEGSSVWVGGRGKYV-L-FDPEGKNNSV-RTVNIQSTKSGCL--DKR 124

```

```

Db      124 edechmfqllalvmsnhlltcgtftafdkpcyidvssfqgverlesgrgkcpfepaqrs 183
QY      125 -D-CENVITLLERSSE-LIACGTNARHPSCNWL-YNG--TYVPLGEMGAYAFSPDSENS 178
Db      184 aavmagvlytavk-nfigteplisravgaedwrlrtetlswlnapfvvaamvlspe 242
QY      179 LVLFEDDEVYSTIRKOEYNGKIPFRFRIRIGESSELYT-SDTV--MÖNQFIKATIV-H-Q 232
Db      243 wgedgddeffftctsvlsyzerikvprvarvcaaglggrkltq-grwtflkadll 301
QY      233 --DQAVDDKITYFFFRDNDPNKNEAPLVNSVAQICRGDGGESSLSVSKMNTFLKAMLV 290
Db      302 cpgpehgrsrylgmaelirpaga-gtrpifgifsqvegaalsavcafrpqdlravln 360
QY      291 CSDAANKNFNRLQDVFLPDPSPGQMDTRRYGVFSNPNMNT---SAYCVYSLGIDKVF-R 347
Db      361 gpfrelkhdcnrglpymdnevpqprgpecianmklqgfsgslspdrvtlfrdhplmd 420
QY      348 TS-SLK-GYHSSLP---N---PRGKCLPQQPIR-TETFOVADRNPE-V-AQR-VE- 392
Db      421 rpvfadgrpllvctdaylvvahrvtslsgkeydvlylgedhlnravrlgag-ls- 478
QY      393 -PMGPKLT-PLEHSK-YHYOKVAVNHRQASHGETFHVLYLTDRGTHKVVPEGEQHSF 449
Db      479 vledlal--fpepqvsesmklyhdw--llvghchevtynsnogrlds-cseclaqdp 533
QY      450 AFNIEIOPFRRAAIQTMSTDAERKLYVSSQWEVSQVPLDLCEVYGGCGHGLMSRD 509
Db      534 vcawfrldacvahahehrgmvgdlesadvsllcpkegpehvvfeyvatvghvlpjs 593
QY      510 YCGMD-Q-GRCISITYSSERSVLQGINPAEPHKECPNRPDAPLQKVS LAPNSRYVLSCP 567
Db      594 psawascvwhqpsgy 609
QY      568 MESRHATYSMRKENV 583

```

```

RESULT 7
ID W17657 standard; Protein: 862 AA.
AC W17657;
DT 24-JUL-1997 (first entry)
DE Human CD100 antigen.
KW CD100 antigen; semaphorin; leukocyte; B cell; T cell; lymphocyte;
KW vaccine.
OS Homo sapiens.
FH Key peptide.
FT Location/Qualifiers
FT 1..41
FT /label= sig_peptide
FT 42..862
FT /label= Mat_protein
FT 42..553
FT /label= Semaphorin_domain
FT 554..630
FT /label= Ig-like_domain
FT 631..733
FT /label= stalk_domain
FT 735..752
FT /label= Transmembrane_domain
FT 753..862
FT /label= Cytoplasmic_domain
FT modified_site
FT 808..815
FT /label= Phosphorylation
FT /note= "putative tyrosine phosphorylation site"
PN M09717368-A1.
PD 15-MAY-1997.
PF 12-NOV-1996; U18645.
PR 09-NOV-1995; US-556422.
PA (DAND ) DANA FARRER CANCER INST.
PI Boussoictis V Freeman GJ, Hall KT, Nadler LM, Schultze JL;
DR WPI: 97-280982/25.
DR N-PSDB: T60655.
PT Nucleic acid molecule encoding CD100 antigen - which stimulates
leukocyte response, e.g. B cell aggregation, differentiation,

```


| | | | | |
|-----------------------|--------|---------------------|-----------------|--------------------|
| Query Match | 7.9% | Score 390; | DB 33; | Length 974; |
| Best Local Similarity | 29.8%; | Pred. No. 1.68e-23; | | |
| Matches | 136; | Conservative 95; | Mismatches 174; | Indels 52; Gaps 36 |

```

Db 113 kdechmfkvl1kknndaellfvcgtnafnpsercnykmdllepfdgfesgmareopydakhan 172
Qy 123 KRDEMYIT-LLERSEBELLACGMNHPSCMNLVNGTVPBLG-EMRGYA--PFSDEPNS 178
Db 173 valfadgklysaetv-dfladavayrsl-geosp-ellrvkhdsakmlkperfyvgaudv 228
Qy 179 LVLEEGDEVEYTIKQOEING-KIRFRFRIGRESLEYTSDIVMNGPQIKATIVHOGQAVD 237
Db 229 dylvffirelaevayltmgkvvfprvaygckindngsgsgrvlekqwtslfkarrlncsvpgds 288
Qy 238 DKRYFFREEDBNPKNPEAPLNVWSRYAOLCRDQGEESLVSXKNFPLKMLVCSAATN 297
Db 289 hfyfnllavevdcvrlrlnp--dv-vlatfstrpynslpgsavaaymdlasiavtgrfkeq 345
Qy 298 K-NENRLODVLLPDPSCQMDRTRYGVFSNPMN-Y--SAYCVYSLDDIDKVF--RTSSL 351
Db 346 kspdstwprvdpdevvpkrpccagsssleryatsnefpddtlnflknp1mdeavpsif 405
Qy 352 KGYHSS---LBNPRPGKLPD--QOPIPTETFFVADHPEVA-Q--RVEP-M--G-P-L- 397
Db 406 nnpvflrmvyrlltklavdcaapy-qnhvtrvlgsekglllkfjarisngfmdslf 464
Qy 398 KTPLE-HS-KYHOKAVAHNPMQSHGTFHVLVLTITDKGRIHKVPEP-GEQE--HSFAF 451
Db 465 leemsvysekcsdygdvedkrlmqdshdtrassllyvaftscvklvplgrcoerhvk-cktk 523
Qy 452 -NIMEI-QPFRRA-AAIQ--TMSLDAER--RKLYVSSQMEVSGVPLDLCEVYGGCGH-G 502
Db 524 ciastdygcylkeggagcshspnsrltftgedlarn 560
Qy 503 GIMSDDPYCGW--DOGRCISLTSSESRVL-OSINPAE 536

```

RESULT 12
ID R71384; standard; Protein; 712 AA.
AC R71384;
DT 21-NOV-1995 (first entry)
DE Tribolium semaphorin I protein.
KW semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW varicella major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
OS Tribolium sp.
PN W0507706-A.
PD 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (RESC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
PI WPI; 95-131177/17.
DR N-PDSB; Q87446.
PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
PS Example 2: Page 85-88; 101pp; English.
CC The sequence of the beetle Tribolium semaphorin I protein. The gene was
CC isolated by PCR using Tribolium genomic DNA. The proteins encoded by the
CC grasshopper semaphorin I (Q87441), human semaphorin III (Q87442),
CC vaccinia virus semaphorin IV (Q87443), Drosophila semaphorin I and II
CC (Q87444-5), Tribolium semaphorin I or varicella major (smallpox) virus
CC semaphorin IV (Q87447) genes were used to generate a series of peptides
CC (R70370-R70418), which retain semaphorin receptor binding activity. The
CC semaphorin derived or semaphorin receptor derived peptides are potent
CC modulators of nerve cell growth, immune responsiveness and viral
CC pathogenesis. They can be used in diagnosis and treatment of neurological
CC disease and neuro-regeneration, immune modulation and diagnosis and
CC treatment of viral and oncological infection and diseases.
SQ Sequence 712 AA;

| | | | | |
|-----------------------|--------|---------------------|-----------------|--------------------|
| Query Match | 7.68; | Score 378; | DB 13; | Length 712; |
| Best Local Similarity | 27.08; | Pred. No. 2.04e-22; | | |
| Matches | 140; | Conservative 128; | Mismatches 194; | Indels 56; Gaps 43 |

[illegible]

RESULT 13 standard; Protein: 650 AA.

ID R1382;
AC R1382;
DT 21-NOV-1995 (first entry)
DE Drosophila semaphorin I protein.
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW varicella major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.

OS Drosophila, sp.
PN W0950706-A.
PN 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (RECC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
DR WPI: 95-131177/17.
DR N-PSDB: 087444.

PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration

PS Example 2; Page 74-77; 101pp; English.

CC The sequence of the Drosophila semaphorin I protein. The gene was
CC isolated by PCR using primers based on sequence homology between the
CC complete grasshopper and partial Tribolium and Manduca (moth) semaphorin
CC gene sequences. The products gave two different sequences, each of which
CC was used to obtain its respective complete sequence: semaphorin I and II
CC (087445). The proteins encoded by the grasshopper semaphorin I (087441),
CC human semaphorin III (087442), vaccinia virus semaphorin IV, Drosophila
CC semaphorin I and II, Tribolium semaphorin I (087446) or varicella major
CC (smallpox) virus semaphorin IV (087447) genes were used to generate a
CC series of peptides (R70370-R70418), which retain semaphorin receptor

CC binding activity. The semaphorin derived or semaphorin receptor derived
CC peptides are potent modulators of nerve cell growth, immune
CC responsiveness and vital pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
SQ Sequence 650 AA:

Query Match 7.4% Score 367; DB 13; Length 650;
Best Local Similarity 28.4%; Pred. No. 2,00e-21;
Matches 124; Conservative 104; Mismatches 158; Indels 51; Gaps 38;

```
Db 3 dcdgnylrmvpspgqlfvcgtnsfmcmtylidsnlyleatkgagcypdphnst 62
Oy 125 DCENYITLLERSEGL-LACGTNARHPSCNN-LVNGTVPLGEMR-GYA--PSPDENS 179
Db 63 svladnelysgtv-adfssdpllyreplqtegydals-lnapny-ssft-qgd-f--- 114
Oy 180 VLEEGDEVYSTRKOEYNGKIPFRRIKGESELYTSDTWQNPOFKATIVHDDQAYDK 239
Db 115 vyffretavefingckaly-srvarvckwdkggphfr-nrwtstlksrlncsfpgdyp 172
Oy 240 IYFFREDNPD-KNEPAPLNVSRVAQLCRGDGSESSLSVSKWNTFLKMLVCSDAATNK 298
Db 173 fyfneiqsasnlveggygmsksllgyvfnctpsnspgsavcafaiaqdadtegfekq 232
Oy 299 -NFNRLQDVF-LLPDSGGMWRDTRVYGVSNNPMN-Y--SAVCYSLGDI-DKV--FR-- 347
Db 233 tglinsnwlpvnmakvpdpypgscndsrldpdtlnflkthlmdenypafsfpllyrt 292
Oy 348 TS-S--LKGYSLSLPNPRPGKCLPDQOPIPETFOVADRHEVAQRVEPMPLTPLFNH 403
Db 293 stlyfctglavdaqlktpgsklydvlfygdthqkllksnaesadsadvtvsvleaidv 352
Oy 404 SK-YHYQKVAHV-RMQASIGETFWLYLTDDKGTTHK-V-VEPGE-QEH--S-FA-FNI 453
Db 333 ltksepitnlelvrtmgydqpkdgsyddgkllivdsqvaiaqprhcnndktitsecvca 412
Oy 454 M-EIQPFRRAAAIQTMSLD-A-E-R--R-KLYVSSQWEVSQVPLDLCEYVG-GGCHGLM 505
Db 413 lqdyrcawdklagkrcs 429
Oy 506 SRDPYCGWDO--GRCIS 520
```

RESULT 14
ID R71383 standard; Protein: 724 AA.
AC R71383;

DE 21-NOV-1995 (first entry)
DT Drosophila semaphorin II protein.
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW varicella major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
OS Drosophila sp.
PN MO9507706-A.
PD 23-MAR-1995.
PF 13-SEP-1994; U10151.
PR 13-SEP-1993; US-121713.
PA (REBC) UNIV CALIFORNIA.
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
DR N-PSDB: 087445.
DR N-PSDB: 087445.
PT New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
PS Example 2; Page 79-82; 10pp; English.
CC The sequence of the Drosophila semaphorin II protein. The gene was
CC isolated by PCR using primers based on sequence homology between the
CC complete grasshopper and partial Tribolium and Manduca (moth) semaphorin
CC gene sequences. The products gave two different sequences, each of which
CC was used to obtain its respective complete sequence; semaphorin I
CC (087444) and II. The proteins encoded by the grasshopper semaphorin I

CC (087441), human semaphorin III (087442), vaccinia virus semaphorin IV,
CC Drosophila semaphorin I and II, Tribolium semaphorin I (087446) or
CC varicella major (smallpox) virus semaphorin IV (087447) genes were used to
CC generate a series of peptides (R70370-R70418), which retain semaphorin
CC receptor binding activity. The semaphorin derived or semaphorin receptor
CC derived peptides are potent modulators of nerve cell growth, immune
CC responsiveness and vital pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
SQ Sequence 724 AA:

Query Match 7.3% Score 362; DB 13; Length 724;
Best Local Similarity 29.0%; Pred. No. 5.63e-21;
Matches 84; Conservative 68; Mismatches 113; Indels 25; Gaps 19;

```
Db 265 vyffretavefingckavy-srvarvckkdvqknlia-hmwaelylarlncsisgefp 322
Oy 240 IYFFREDNPD-KNEPAPLNVSRVAQLCRGDGSESSLSVSKWNTFLKMLVCSDAATNK 298
Db 323 fyfneiqsavyqpsdksrffat-ftstnglgsavcsfhinegaafngkfkqsssn 380
Oy 299 -NFNRLQDVFLLPDSGGMWRDTRVYGVSNNPMNSAVCVSLGDDI--KV-FR--RS- 349
Db 381 sawlpvlnsrpeprgctcvndtsnlpdtvlnflrshplmdkavnh-eh-npyykrdl 438
Oy 350 S--LKGYSLSLPNPRPGKCLPDQOPIPETFOVADRHEVAQRVEPMPLTPLFNH-X 406
Db 439 vftklvavkklidllngylyyyvgtnlgrlyklyqyrngesls-klldifevapneal 497
Oy 407 HYQKVAHRMQAS-HGEFFHWLYLTDDKGTTHKVAPEDESHSEAFNIMEIQPFRRAAAI 465
Db 498 qymeisqrkxlylqtdhrklyqldiamcnrrydh-cfrc-v-rdpycgwd 544
Oy 466 QTMSLDARRRLIYVSSQWEVSQVPLDLCEV-YGGCHGLMSRDPYCGMD 514
```

RESULT 15
ID W57260 standard; Protein: 930 AA.
AC W57260;

DT 02-SEP-1998 (first entry)
DE Human semaphorin Y.
KW Human; semaphorin Y; nerve extension inhibitor; central nervous system;
KW peripheral nerve growth.
OS Homo sapiens.
PN MO9811216-A1.
PD 19-MAR-1998.
PF 09-SEP-1997; J03167.
PR 08-AUG-1997; JP-227220.
PR 11-SEP-1996; JP-263565.
PA (SUMU) SUMITOMO PHARM CO LTD.
PI Kikuchi K, Kimura T;
DR WPI: 98-250958/22.
DR N-PSDB: V28915, V28916.
PT DNA encoding human and rat semaphorin Y - an inhibitor of nerve
PT extension
PS Claim 1; Page 65-70; 85pp; Japanese.
CC The present sequence represents human semaphorin Y which inhibits
CC nerve extension. Semaphorin Y genes and proteins may be used to inhibit
CC peripheral nerve growth. Semaphorin Y antagonists can be used to
CC accelerate regrowth of the central nervous system.
SQ Sequence 930 AA:

Query Match 7.1% Score 352; DB 32; Length 930;
Best Local Similarity 28.4%; Pred. No. 4.45e-20;
Matches 132; Conservative 112; Mismatches 164; Indels 56; Gaps 44;

```
Db 120 ecnylrvlvpwsgqlacgtnsfpcrsgytlsgqgeglsgqarcpdatgsnva 179
Oy 125 DCENYITLLERR-SEGLACGTNARHPSCNNLVNGTVPLGE-MRGYA--PSPDENSIV 180
Db 180 lfegeilyata-adfqsadavvyslqpppl-ts-a-kydskwlrephfvgalehgdh 235
```



```
OY 181 LFBGEVYSTIRKOEYNGKIPR-FRRIRGESELYTSPTVMQNPQIKATIVHQDQAYDK 239
Db 236 vyfffevsvedarlgvqfsvrvarvckrdmggsprald-rhwtsfklrlncs-vpgds 293
OY 240 IYFFREDNDPKNEAPLUNSRVAQLCRGDOGSES-SLSVSKWTFPLKAMLYCSDAATNK 298
Db 294 tly-f-dvlgalltpvnlhgrsa-lfgvftqlnslpgsavcafyldelergfegkfkq 350
OY 299 NFNRLODVFL-LPDPSC-QWMDTRVYGVFSNPN-Y--SAVCVSLGDIDKF--R-TS- 349
Db 351 rsldgawtpvseavpseprpgscagvgaaalfssralpddvltfikhplldpavpvt 410
OY 350 -SLKG-Y-H-S-S-LPNRPREKCLP-DQOPIPRETFQVADRHPE-V-AQRY-EP-MGPLK 398
Db 411 hqpllttsralltqvavdgmagphsn-itymflgsndgtvklvtggrsggpepllle 469
OY 399 -TPLFH--SKYHOKVAVHMQASHGETFHYLYLTDRGTIHKVYEPGEOEHSAFNIME 455
Db 470 eldaysparcsqkrtaqtarrilglldteghrlfvaifsgciyyllpiscarhga-cqrs 528
OY 456 -IQPF--R---RAA--A--IQTMSLDAERRKLYSSQMEVSGVPLDLCEVYGGCH-G 502
Db 529 clasqdbpcgwhsrgcvdlrsgsgtdvdqagngesmehgdcq 572
OY 503 CLMSRDEPCGMDQGR-CISYSSSE-RSVLOSINPAE-PHKECPN 543
```

Search completed: Tue Aug 3 16:26:35 1999
Job time : 134 secs.

This Page Blank (uspto)

| | | | |
|----|-----|--|-----|
| Db | 301 | NRLDVLFLDPBSQWMDTVYGVENSPNNYSACVYSLIGDIDKFKRSSLKGYHSSLPN | 360 |
| Qy | 301 | NRLDVLFLDPBSQWMDTVYGVENSPNNYSACVYSLIGDIDKFKRSSLKGYHSSLPN | 360 |
| Db | 361 | PRPGKCLPDDOPIPTETFFQVADRHPEVAORVEMPGRLKPLFHSYHYHOKAVYRMASH | 420 |
| Qy | 361 | PRPGKCLPDDOPIPTETFFQVADRHPEVAORVEMPGRLKPLFHSYHYHOKAVYRMASH | 420 |
| Db | 421 | GEIFHVLYLTDTKGTIHKVVEPGEQESHFAFIMEIOFERRAALIQTMSIDAERKKLYVS | 480 |
| Qy | 421 | GEIFHVLYLTDTKGTIHKVVEPGEQESHFAFIMEIOFERRAALIQTMSIDAERKKLYVS | 480 |
| Db | 481 | SQWESVQVPLDLEVYGGGCGHGLMSRDYCGMDGRCISYSSRSVLOSINPAEPHKE | 540 |
| Qy | 481 | SQWESVQVPLDLEVYGGGCGHGLMSRDYCGMDGRCISYSSRSVLOSINPAEPHKE | 540 |
| Db | 541 | CPNPKPKAPLOKVS LAPNSRYTLCSPMSRATYSWRHKEVBOCSCEPGHOSNCILFI | 600 |
| Qy | 541 | CPNPKPKAPLOKVS LAPNSRYTLCSPMSRATYSWRHKEVBOCSCEPGHOSNCILFI | 600 |
| Db | 601 | ENLTAQOYGYHYFCEAOGSYFREDQHMQLLPEDGIMAEHLGLHACALAASTLMLGVPTLT | 660 |
| Qy | 601 | ENLTAQOYGYHYFCEAOGSYFREDQHMQLLPEDGIMAEHLGLHACALAASTLMLGVPTLT | 660 |
| Db | 661 | LGILLVH 666 | |
| Qy | 661 | LGILLVH 666 | |

| | | | | |
|-------|---|--|----------------|------------|
| | RESULT | 2 | | |
| ID | 088371 | PRELIMINARY; | PRI; | 393 AA. |
| AC | 088371; | | | |
| DT | 01-NOV-1998 (TREMBLREL_08 | CREATED) | | |
| DT | 01-NOV-1998 (TREMBLREL_08, | LAST SEQUENCE UPDATE) | | |
| DT | 01-NOV-1998 (TREMBLREL_08, | LAST ANNOTATION UPDATE) | | |
| DE | SEMAPHORIN L (FRAGMENT). | | | |
| GN | SEMAL. | | | |
| OS | MUS MUSCULUS (MOUSE). | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA; | | | |
| OC | SCUROGNATHI; MORIDAE; MORINAE; MUS. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE: 98389619. | | | |
| RA | LANGF C., LIEHR T., GOEN M., GEBHART E., FLECKENSTEIN B., ENSSER A.; | | | |
| RT | "New eukaryotic semaphorins with close homology to semaphorins of DNA | | | |
| RL | viruses". | | | |
| EMBL: | AF030699; G3523117; -. | | | |
| DR | NON_TER | 393 | | |
| FT | SEQUENCE | 393 AA; 43909 MW; 7CB8193C CRC32; | | |
| SO | | | | |
| | Query Match | 52.2%; Score 2595; DB 11; Length 393; | | |
| | Best Local Similarity | 90.2%; Pred. No. 0.00e+00; | | |
| | Matches | 357; Conservative | 18; Mismatches | 17; Indels |
| | | | 4; Gaps | 2; |
| Dh | 1 | MTPPPPGRAPASARAVVLSPARFGGLPLRLILLVVAAASAQGSBSGPRISAWKG | 60 | |
| Oy | 1 | MTPPPPGRAPASARARVPGPAPRLGLPLRLLLLLMAAASQGILRSGPSPIFAWKG | 60 | |
| Dh | 61 | --ODHWDFSQPEPHIVLFHEPGSFYSVWGVRGVYHFNFPEEGKNASVTRVINIGSTGSG | 117 | |
| Oy | 61 | HVGGDVRDVFEGTETHYLVLFHEPGSSSVWVGRCGVLYLFDEPDECKMASVRVINIGSTGSG | 120 | |
| Dh | 118 | QDKODCCNGYTTLERBNGLLVCGTNAKRPSCMNLVNDVSVMSLGEMKGYAPFPDENSL | 177 | |
| Oy | 121 | LDRDCDENYTTLERREBGLAGCTNARNRSHSCMWLVNGVTYV-ELGEKRGYAPFPDENSL | 179 | |
| Dh | 178 | VLEFGDEVYSTIRKOENGKIIPRRRIARGSELYSTPTWMONOGFITATVIHDDOAYDDK | 237 | |
| Oy | 180 | VLEFGDEVYSTIRKOEINGKITPPRRRIARGSELXTSDPTWMONOFITATVIHDDOAYDDK | 239 | |
| Dh | 238 | IYYFFREDNDKNKEADLANVSRVAQLCRGDGGESSLSVKWNTPFLKAMIVCSDAATNRN | 297 | |

[illegible]

| RESULT | 3 | PRELIMINARY; | PRT; | 653 AA. |
|--------|---|---|------|---------|
| ID | 064906 | | | |
| AC | 064906: | | | |
| DT | 01-NOV-1996 | (TREMBLREL. 01, CREATED) | | |
| DT | 01-NOV-1996 | (TREMBLREL. 01, LAST SEQUENCE UPDATE) | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | |
| DE | SIMILAR TO GENBANK ACCESSION NUMBER L26081. | | | |
| OS | ALCELAAPHINE HERPESVIRUS 1. | | | |
| OC | VIRUSES: DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE; | | | |
| OC | GAMMAHERPESVIRINAE. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97201573. | | | |
| RA | ENSSER A., FLECKENSTEIN B.; | | | |
| RT | "Alcelaphine herpesvirus type 1 has a semaphorin-like gene."; | | | |
| RL | J. GEN. VIROL. 76:1063-1067(1995). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | ENSSER A., PFLANZ R., FLECKENSTEIN B.; | | | |
| RT | "Primary structure of the alcelaphine herpesvirus 1 genome."; | | | |
| RL | J. VIROL. 71:6517-6525(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-C500; | | | |
| RX | MEDLINE; 97404659. | | | |
| RA | | | | |

| | Query Match | 38.4% | Score 1907 | DB 14 | Length 653 |
|----|--|--------------|--------------------|----------------|------------|
| | Best Local Similarity | 46.2% | Pred. No. 0.00e+00 | | |
| | Matches 279 | Conservative | 108 | Mismatches 203 | Indels 14 |
| | | | | | Gaps 14 |
| Db | 45 PAAMGILCVSIRLMLL-SAITAKSRFDIKRLIYNLTDFGQHRF-FGPOEPHTVLFH | 102 | | | |
| Qy | 22 PARLG-LPRLRLRLLLMMAAAGQHLNSGRIFVWKGHNGQDRYDGGQLEPHRTVLFH | 80 | | | |
| Db | 103 SLNSDDVYVGNNNTYLFDFAHSSNASTALINTSTHNTHRLSSTCFENFTLLHQDTGL | 162 | | | |
| Qy | 81 EPSSSSVWVGGRKYLTFDFPEGKNASVYTVNIGSTKSGCLDKRCDCENYITLLERRSGL | 140 | | | |
| Db | 163 LAGTNSOKPSCW-LINNLDTQELGPKLCLAPSPSSGULVFDONDYISTNLKSLSG | 221 | | | |
| Qy | 141 LAGTNAHRSPSCNWLNVGVVP-LGEMRAYFPSPDENVLFEDEDEVYSTRK-QEYNG | 198 | | | |
| Db | 222 S-EKFRRIAGOVELYSDTAMHRPOCVATVAHKNSYDOKLYPFOENSHSDFKOPHT | 280 | | | |
| Qy | 199 KIRFRFRIGESLTYSDTVMQNPQRIKATITYHOADAIDDKIYTYFREDNPDKNEAPLN | 258 | | | |
| Db | 281 VPRVGVGSSDDGGESSLVSVKMTFLKARLACVDYDGRINELQDIFITWAPENSWE | 340 | | | |
| Qy | 259 VSRVAOLCKGDGGESSLVSVMKMTFLKMLVCSDAATKKNENRQDYFLLPDEGGQRD | 318 | | | |
| Db | 341 TLTYGLFSPNPFNSANCYVTVKDIDHVFYTSKLYNTHHKLPPRPGCKMKNHQVPTET | 400 | | | |
| Qy | 319 TRYVGVFSPNPNVNSACVYSLDGDIVKFTSTLKGHSLPNPRGKCLPDDQIPPTET | 378 | | | |

```
Db 401 QVADRYEVADPVYQKNNNAMEPIIQSKYITKLLVYREYQ-GVFWATIFLYTIKGTIH 459
      ||||| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
      379 QVADRHPEVACRVEMGFLKTPLEFSKYHQKAVHMQASHGETHF-VLYLTDRGTIH 437
      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
      460 IYVREDSNSTALNILEINFQKAPQIQLDNTNKLKLVNSWEVSEVPLDLSYVG 519
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      438 KVPVEGEDESHSFANIMEIQPFRRAMAQIOTMSLDAERKRLVSSQWVSVQVPLDCEYVG 497
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      520 NDSCSEMSRDLCTWYNNTC-S-FK-GRVSVEGPGANRTLSEMGGBHYAPTYKHQYS 576
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      498 GCGHGLMSRDPYCGWDGRCISTYSSRSVLSQINPAEPH-KE-CPNPKPKAPLQKVS 555
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      577 IPLNSVSLSCPAPVSNHADYFWTKDGFKEKCHVYTHKNCICILLANSTTATNTHVQNM 636
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      556 LAPSRVYLSLCPMESRHAITYSMRKHEVQSCPEGHOSPNCILFENLTAAQYGHYCEA 615
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      637 KEDS 640
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
      616 QEGS 619

RESULT 4
ID 062177 PRELIMINARY; PRT: 748 AA.
AC 062177;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN A PRECURSOR (SEM A).
GN SEMA OR SEMA.
OS MUS MUSCULUS (MOUSE).
OC EUKAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCURGNATHI; MURIAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE: 95267431.
RA PUSCHEL A.W., ADAMS R.H., BETZ H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL NEURON 14:941-948(1995).
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR (PROBABLE).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO.
CC LOW LEVELS FOUND BETWEEN DAYS 10-12.
CC EXPRESSION PEAKS ON DAY 13 WITH MODERATE LEVELS FROM THEN UNTIL
BIRTH.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC CONTAINS ONE C2-LIKE DOMAIN.
DR EMBL: X85990; G854324; -.
DR MGD: MGI:107561; SEMA.
DR PFM: PFO0047; 1g; 1.
KW SIGNAL; IMMUNOGLOBULIN FOLD; MULTIGENE FAMILY; NEUROGENESIS;
KW DEVELOPMENTAL PROTEIN.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 748 SEMAPHORIN A.
FT DOMAIN 586 649 IG-LIKE C2-TYPE DOMAIN.
SQ SEQUENCE 748 AA; 82894 MW; A7E53A8D CRC32;

Query Match 14.8%; Score 737; DB 11; Length 748;
Best Local Similarity 33.7%; Pred. No. 1,22e-150;
Matches 168; Conservative 112; Mismatches 166; Indels 52; Gaps 36;

Db 112 ECNNFRLHAYNHTLIACRTGAHPTCALMRNATAGTHASTGPEKLEDKGKTPYP 171
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      125 DCENVITLLERRSDEG-LIACGTNARHPSC---NWLVNGT-V-V-P--LGEMRGVAPRSP 174
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      172 RHRPSVYVGEELYSGV-TAALMGDRFTIFRSISLGNPSLAFEPHRSRLNPKPKFYKRWI 230
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      175 DENSLVLFEGDEVYSTIRKQYNGK-IPRRIRIGESSELYT-S-DTV-MONPOFIK-I 229
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      231 PESEMPDDDKTYFFRERSAVEAAPAMGRMSVSRVQICRNDLGGORSL-VNKKWTFELKAR 289
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
```

```
QY 230 VHDQAVADKIYFFREDNPNKDE-APLVNSRVAQLCRDQGGESSLSVKMNTFLKAM 288
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      290 LVCSVPGEQDTHFDQLODFLLS-SR-DROTPLLAVFSTSSGVFOGSAVCYSMNDVR 347
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      289 LVCS--DAATKKNRNLQDFVLLPDPGQMRDITVYGFSPW-N-Y--SANCYSLGDDID 343
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      348 RAFLGPLPHKGPYHOMVYSGRVYRPGMCPSPKTEGTSSTKDFPDVYIQGRNHPLM 407
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      344 KVERTS-SLK-G-----YHSSLPNRPGKCLDPQ-QPIR-TETF--QVAD--R-HREV 387
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      408 YNPVPMGG-R-PLFLOYGAGYTTQTAAADVAAADGH-IDVLEIGDVGIVLVISVPK 464
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      388 AQREPMGPRLKTPLE-H-SKYHYQKVAVHRMQASHGETFHVLYLTDRGTIHKVE-P- 442
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      465 GRNRSGLLLELOVEDSALITSMQISSKROOLYVASRAVAOIALHRTALGRACAE 524
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      443 GEOHSEFANIMEIQPFRRAMAQIOTMSLDAERKRLVSSQWVSVQVPLDCEYGGCHG 502
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      525 CCLARDPYCAMPDSACTRFOPTAKRRFRRODIRNGDSTLCSGDSHVLLEKTVLGVES 584
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      503 CLMSRDPYCGWDGRCISTYSSRSVL--QSINPAEPHKECPNPKPKAPLQKVS LAPNS 560
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      585 GSAPLECEPRSLQAHVQM 602
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      561 -RYVLSCPMESRHAITYSM 577

RESULT 5
ID 062181 PRELIMINARY; PRT: 751 AA.
AC 062181;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN E PRECURSOR (SEM E).
GN SEMAE OR SEME.
OS MUS MUSCULUS (MOUSE).
OC EUKAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCURGNATHI; MURIAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE: 95267431.
RA PUSCHEL A.W., ADAMS R.H., BETZ H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL NEURON 14:941-948(1995).
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR (PROBABLE).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO.
CC MAXIMUM EXPRESSION BETWEEN DAYS 10-12 WITH MODERATE LEVELS FROM
DAY 13 UNTIL BIRTH.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC CONTAINS ONE C2-LIKE DOMAIN.
DR EMBL: X85994; G854332; -.
DR MGD: MGI:107557; SEMAE.
DR PFM: PFO0047; 1g; 1.
KW SIGNAL; IMMUNOGLOBULIN FOLD; MULTIGENE FAMILY; NEUROGENESIS;
KW DEVELOPMENTAL PROTEIN.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 751 SEMAPHORIN E.
FT DOMAIN 587 649 IG-LIKE C2-TYPE DOMAIN.
SQ SEQUENCE 751 AA; 85259 MW; B28D6CFE CRC32;

Query Match 14.6%; Score 725; DB 11; Length 751;
Best Local Similarity 33.0%; Pred. No. 1.53e-147;
Matches 148; Conservative 110; Mismatches 149; Indels 41; Gaps 28;

Db 167 SFNPNVNTVSMINEELPFSGNT-DFMGTDAIRSLTKRQLTDOHNSKWLSEPMFVD 225
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      171 PFSDENSLVLFEGDEVYSTIRKQYNGK-IPRRIRIGESSELYT-S-DTV-MONPOFIK 226
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
      226 AHVDPGDPAADVAVYFFERKRLTDNNRSTKQIHSMTARICPNDTGGORSL-VNKKWTFEL 284
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
```

[illegible]

```

RESULT      6
ID          013214      PRELIMINARY;      PRT:      749 AA.
AC          013214;
DT          01-NOV-1996 (TREMBLREL, 01, CREATED)
DT          01-NOV-1996 (TREMBLREL, 01, LAST SEQUENCE UPDATE)
DT          01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)
DE          SEMAPHORIN V.
OS          HOMO SAPIENS (HUMAN).
OC          EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES
OC          CAVIARHINI; HOMINIDAE; HOMO.
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE: 96210603.
RA          SKRIO V., BADER S., LATIF F., CHEN J.Y., DUH F.M., WEI M.H.,
RA          ALBANDSI J.P., LEE C.C., LEMMAN M.I., MINNA J.D.;
RT          "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT          cancer deletion region and demonstrate distinct expression
RT          patterns.";
RL          PROC. NATL. ACAD. SCI. U.S.A. 93:4120-4125(1996).
DR          ENBL: U28369; G974284; -.
SQ          PRAM: PF00047; I9: 1
SQ          SEQUENCE 749 AA: 83121 MW; 6138C5D5 CRC32;

```

| | 14.4%: | Score | 716: | DB | 4: | Length | 749: |
|-------------|--------------|--|------------|------|-----------|--------|------|
| Query Match | Similarity | 33.0%: | Pred. | No. | 3.2E+145: | | |
| Best Local | Conservative | 115: | Mismatches | 170: | Indels | 54: | Gaps |
| Matches | 167: | | | | | | |
| D b | 112 | ECCMFVKLIHAAYNRTHLACGTGAFHPYCAVEVGHRAEEFVLRLDPGRIDEGKGSYPD | 171 | : | : : | | : : |
| Q y | 125 | DCENTITLLER-RSEGLIACGTNMRHPSG-WNLV-N-G--IVYPL--G--EM-RGYAPFS | 173 | : | : : | | : : |
| D b | 172 | PRHRAASLVGEELSGYA-ADLMGRDFTIRRSIGQRSLTEPHDSRWLNLEPAVFVFW | 230 | : | : : | | : : |
| Q y | 174 | PDENSLVFEEDEVYSTRKOEYNCK-LPRFRRIERGESELYT-S-DWV-MONPOFIKAT- | 228 | : | : : | | : : |
| D b | 231 | IPESENPPDDKIYFEFRETVAEAPAALGRLSVSVGQICRNDVGGKSL-NKKTKTEFLKA | 289 | : | : : | | : : |
| Q y | 229 | IYHODADAIDDKIYIFRRDNPKNPE-APLVASRAOQCRDGOGESLSKVNITLKA | 287 | : | : : | | : : |
| D b | 290 | RLYGVSPVEGDTHFDLODYFLLS-SKH-D-RIPLLYAVETSSSIFGSAVCYSMNDV | 347 | : | : : | | : : |
| Q y | 288 | MLVCS--DAATNNKNFNRLQDFVLLPDPGQQRDRTRVYGVSFPANN-Y-SAVCYVSLGDI | 342 | : | : : | | : : |
| D b | 348 | RRALFGPAHHDEGPPIHOWSVYQGKRVRYPRPMGCSKITGTSTSKIDPPDYIQARAHPL | 407 | : | : : | | : : |

QY 343 DK-V-F--RTSSL-R-GYHSLPRRRCGLPQ-QIP-TERF--QVAD-R-HHE 386

Db MYNSVLEPTGG-R-PLFIQVCANITFTTOIAADRYAADAGH-YDYLETGTIDVGLKATISVP 464

QY 387 VAOREEMPGRLKPLP-H--SKYHOKVAHHRQASHGEGTFHVLTLTTRGITTHKVE-P 442

Db KGSRPASGELLLEELHAYFEDSAATYSMOISKRKHOLYVAGRSVAADIALHRCAGHRYCT 524

QY 443 -GEDEHSAFNIIMTIOFFRRRAAAIQYTLSLAEKRKLITYSSQMEVSQVPLDLEVTYGGGH 501

Db 525 ECIILARDPYCAMDVGACTRFQPSAKRRFRRODYRNGDPSTLSCGSDSSRPALLEHKYGV 584

QY 502 GCLMSRPDPCGMDGRCISITSSRSYVL-QSINPAEPHKECENPKPRKAPL-QKYSIAP 558

Db 585 GSAFLCECEPRSLQARVETFTFRAGV 610

QY 559 NSRYTLSCPMESRHATYSMR-HKENV 583

```

RESULT 7 PRELIMINARY: PRT: 750 AA.
ID Q93018
AC Q93018;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN V.
OS HOMO SAPIENS (HUMAN).
OC EOKARROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES.
OC CATARRHINI; HOMINIDAE; HOMO.
[1]
RN RP SEQUENCE FROM N.A.
RA DANTE M., WAMBLEY P.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN RP [2]
RL SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN RP [3]
RL SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U73167; G2880035; -.
DR PRAM: PF00047; 19; 1.
Q SEQUENCE 750 AA; 83034 MW; C8C87E34 CRC32;

```

| Query Match | Similarity | Score | DB | Length |
|--------------|-------------------|--|------------|----------|
| Best Local | 33.0% | Pred. No. 3,21e+145; | | |
| Matches 167; | Conservative 115; | Mismatches 170; | Indels 54; | Gaps 41. |
| Db | 113 | ECMNFVKLLHAYNRTLLHACGTGAFHPTCAFEVGHAEHPVLRLLDGRITEDGKSGSPYD | 172 | |
| Qy | 125 | DCENVITLLER-RREGLLACNTAHRHSC-WNLV-N-G--TVPL--G--EM-RGYAPFS | 173 | |
| Db | 173 | PRHRAASVLVGEELYSQVA-ADLMGRDFTTFRSLGQRPSSLRTPEHDSRWLNPEKVFVFW | 231 | |
| Qy | 174 | PDENSLVLFEEDEYVSTIRKQYNGK-IPFRFRIRGESELYT-S-DTV-MQNPFIKAT- | 228 | |
| Db | 232 | IPESBNDDDKIYFFFEETAVEAPALGRISLVSVQGITCNDVGGQSL-VNKKTTFLKA | 290 | |
| Qy | 229 | IVHODQYDKIKYFFREDNPDKNPE-APLNVSVVALCGDGGESLSVSKWTFLLKA | 287 | |
| Db | 291 | RLVQSVGVEGDDHFFDLOLVFLLS-GRDH-RPFLAEVSTSSIRQGSANVCVSNVDV | 348 | |
| Qy | 288 | MLVVS--DAATNKNFENLQDVFLLPDSGGWRDTRVYGVSNPNP-Y--SACVYISLGD1 | 342 | |
| Db | 349 | RRALGFPAHKEGEMQWVSQGHVPYPRGMCBCKTFTGSSITKDFPDVDVIOFARNHPL | 408 | |
| Qy | 343 | DK--V--F--RTSSL-K--GYHSSLPMRPGRCKLPDQ-QP1P-TETP--QVAD--R-RPE | 386 | |
| Db | 409 | MYNLSVLPTGG-R-PLFLQVGCANTFTIADRAAADGH-YDVLFIQDVGVLKATVSP | 465 | |
| Qy | 387 | VAORVEEMGPKLPPL-H--SKYHOKVAVAHROMASGELFFHVLVLLTLDGTLHKVE-P | 442 | |

Wed Aug 4 09:15:32 1999

US-09-240-410-2.rspt

Page 5

[illegible]

| RESULT | ID | 8 | PRELIMINARY: | PRT: | 751 AA. |
|--------|---|-----------------|-------------------------|------|---------|
| AC | 099985 | | | | |
| DT | 01-MAY-1997 | (TREMBLREL. 03, | CREATED) | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, | LAST SEQUENCE UPDATE) | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, | LAST ANNOTATION UPDATE) | | |
| DE | SEMAPHORIN E. | | | | |
| OS | HOMO SAPIENS (HUMAN). | | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATE | | | | |
| OC | CATARRHINI; HOMINIDAE; HOMO. | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | YAMADA T., ENDO R., GOTOH M., HIROHASHI S.; | | | | |
| RT | "identification of semaphorin E as a non-MDR drug resistance gene of | | | | |
| RT | human cancers." | | | | |
| RL | PROC. NATL.ACAD. SCI. U.S.A. 94:14716-14718(1997). | | | | |
| DR | EMBL: AB000220; D1033360; - | | | | |
| DR | SEQUENCE 751 AA; 85207 MW; F89FB934 CRC32; | | | | |

| | | | | |
|-----------------------|--------|----------------------|--------|-----------------|
| Query Match | 14.38; | Score 710; | DB 4; | Length 751; |
| Best Local Similarity | 33.68; | Pred. No. 1.13e-143; | | |
| Matches | 151; | Conservative | 104; | Mismatches 151; |
| | | | Indels | 43; |
| | | | Gaps | 30 |

| | | | |
|--------|--------------------------------------|--|--------------|
| Db | 167 | SENPMVNVSMINDELFSGMTI--DEMOTDAIRERSLTKRNAVYTDQHNKMWLSPEPHVD | 225 |
| Oy | 171 | FFSPDENSLVLFEGEVEVSTIRKOEYDNK--IPRRRLRGESELVTS--DIV-MONPOFIK | 222 |
| Db | 226 | AHVIDGDDPMDAKVYFFFEKELTDNNSTKQISHMARICPNDTGGRLSL-VKRWTFPL | 284 |
| Oy | 227 | ATIVHOD-QAIDDKIYYFFREDNDPKNEAPLANSRYAQOLCRGOGGESSLSYSKNMTFL | 285 |
| Db | 285 | KARLVCSYTDDEGGEYHDELEDFLL-ETDNP-RTLVYGIFTTSSVFEGSAVCYHL | 342 |
| Oy | 286 | KAMLVCS--DA-ATNKNFNRLQDFVFLBPDPSGQMRDFRVYGVFSNPMW-Y--SAVCYSL | 333 |
| Db | 343 | SDIQTVENGPRAHKEGPNHOLISIOGRIPIYRPBGTCGEGATPMNRTTKKEPDDVYFIR | 402 |
| Oy | 340 | GDIDKVFRTS-SLK-G-----YHSSLPNRPBGKCLPDC-QP-IPET-F--OVAD--R | 383 |
| Db | 403 | NHPLMYNSIYPIHKRPLVRI-GDYKYKTAIVDRVYNAADG-RKHVFLGDRGTQKVY | 466 |
| Oy | 384 | HPEVAQRVEPM-G-PLKTPLEFHSYHAIQKAVHMQOSBETHVLLITLDKRTIRHVY | 444 |
| Db | 461 | VLPTNVSNGEELLELEFVNKNHAPITMTKISSKKOOLYSSNGEVSIVSLHRHIGTA | 520 |
| Oy | 441 | E-PGEQHSFPAFNIMEIOPFRRAALIQMSLDAERKIKYVSSQMEVSOVPLDLCEYVGG | 495 |
| Db | 521 | CADCCCLARDPYCAMDGHSCSFYPTGKRRSRQDVRRHGNPLTJCRGFNLKAYRNAELIVQ | 580 |
| Oy | 500 | CHGCLMSRDPYCGMDOGRGCIISYSE--RSVLQGISNPAEPHKECP--NPKPDAPLQKVS | 555 |
| Db | 581 | YGVKNMTTFLECAPKSPQASIKMLQKKK | 609 |
| Oy | 556 | LAP-NSRYLSCPMESRRAIYSWR-HKEN | 582 |
| RESULT | 9 | | |
| ID | 042236 | PRELIMINARY: | PRT: 751 AA. |
| AC | 042236: | | |
| DT | 01-JAN-1998 (TREMBLERL. 05. CREATED) | | |

DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE COLLAPSEN 3.
 OS GALLUS GALLUS (CHICKEN)
 CC EUDARTIDA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
 CC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA FEINER L., KOPPEL A.M., KOBAYASHI H., RAPER J.A.;
 RL NEURON 15:0-0(1997).
 DR EMBL, AF022946; G2522204; -.
 SQ SEQUENCE 751 AA; 85432 MW; 0B9BA2F7 CXC32;

Query Match 14.2% Score 705; DB 13; Length 751;
Best Local Similarity 33.5% Pred. No. 2,19e-142;
Matches 150; Conservative 107; Mismatches 150; Indels 41; Gaps 28

```
Db 167 SFNPNTVTSVMINEELFSGMYT-DENGTDAIATIRSLTKRNAVTDHNSKMLSEPIFVD 225
QY 171 PPSDENSLTVFEEGDVYSTRKOEYNGK-IPRRRIIGSELYTS--DTV-WONPQIF 226
Db 226 AHHVDPGDNDPAKITYEFFEKRLDNGSGSTKQIHSIARICPDNDGGORSU-VNKKTF 284
QY 227 ATIHQD-QAIDDKIITYFFREDNDKKPEAPLANSRQAOLCRDGGESSLSVSKNMTFL 285
Db 285 KARLVCSYMDGTEYTFDELEDFLL-ETDNP-RITLVYGIFTTSSIFKGSAYCYHL 342
QY 286 KAMLVCS--DA-ATNKNNENRLQDFELLDPDSSGQNRDTRVYGVFSNPNY-Y-SAVCYSL 339
Db 343 SDIQTFVNGPRAHNEGPHQILPQGHILPYRPQETCGGAFTPMRTKTEFPDQVVFIR 402
QY 340 GDIDVFTST-SLK-G-----YHSLPNRPKCLPDO-QP-IPET-F--QVADRRP 385
Db 403 NHPLATNPIYPIHKKRPLIRIGTDYKTKIAVDKNADNRG-HVLELGTDOGTQKRVV 461
QY 386 EVAQREVEMGRL-KTPLF-H--SKYHQVAVAHMQASHGFTFHVLLTDRGTHIKHVE 441
Db 462 LPTNFSASGELLLELEVFQSNSTITTMKISSKKQOLYSSSEVTOVPLRHCIYGTAC 521
QY 442 -PGDEHSEFAFNIMEIOPFRRAALIQMTSDAERKRLYSSQWVSQVPLDLECYVGGC 500
Db 522 ADCCCLAPRYCAMPQSGNSCRPEYPGKKRRRODVRHGNPILQCRGFLMLKAYNAETVOY 581
QY 501 HGCLMSRPYCGMQOGRISTYSSE-RSVQLSTINPAEPHKECP-NPKEDKAPLOXVSL 556
Db 582 GVKNNVTFLECTPKSPQASINMILLKDN 609
QY 557 AP-NSRYLLSCPMESRAHYSMR-HKEN 582

RESULT 10
AC Q13372 PRELIMINARY; PRT; 753 AA.
ID Q13372;
Db 01-NOV-1996 (TREMBLREL_01, CREATED)
Dt 01-NOV-1996 (TREMBLREL_01, LAST SEQUENCE UPDATE)
Dt 01-NOV-1998 (TREMBLREL_08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN III FAMILY HOMOLOG.
DE HOMO SAPIENS (HUMAN).
OC EDUAROTIA, METAALCA: CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE; 96230324.
RA XING R.H., HENSEL C.H., GARCIA D.K., CARLSON H.C., KOK K., DALY M.C.,
RA KERBACHER K., DEN BERG A., VELDHOUS P., BOYS C.H., NAYLOR S.L.;
RA "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
RT 3p21, a region deleted in lung cancer*";
RL GENOMICS 32:39-48(1996).
DR EMBL: U38276; GI061351. -.
PFAM: PF00047; Ig; 1.
SQ SEQUENCE 753 AA; 84941 MW; BECBBB0 CRC32;
```

[illegible][illegible]

```

OY 133 LERRESE-LMACGINARHPS-C-W-NL-VN--GIYVPL--G--EM-RCIAPFSPDENSIVL 181
Db 181 LVDEGLIYSIGTA-ADPMGRDEAIPTFTLGHHPHRTIOHDSFMLNDPRISAHILIPESDNPE 239
OY 182 FEGDEVUSTIRKOEYNGK-IPFRIRIGESIELYTS--DTV-MONPOGIKATIVHO--DAQY 236
Db 240 DDKIYFEFFERENAIOGEHNGKATHARIOICOKNPGGHRSL-VUKMITLEKLALICSVPP 298
OY 237 DDKIYFEFFERENPPKNEPAPLYNSRVAQOLCKRGDGGESSLSYKMWTFKLAMLYTS--DA 294
Db 299 NGIDTHEDELQVFLM-NSKOP-KNPIYGVFTTSSNIEFGSAVCMSYMTDVARVYFLGY 356
OY 295 -ATKNENRLODQVELLDPPSGOMHDITRYGVGVFSNPNM-X--SAVCYSLGIDDKVERIS- 349
Db 357 AHRGPNQWVRYGGRVRYPRGCSBKITGGSDRSTKOLRDEYITFARSHPMYNPVPI 416
OY 350 SLK-G-YH-S-S-LPNRPCKCLDDQ-OPJ-PIETP--QY-A-DR-HPEAORVERME 394
Db 417 NS-RPIMIKTVDVQFTQIVVDRKDAEDGQ-YUMVEIGTDIGTVLKVSIIPKETWHELE 474
OY 395 GPLTLPFLHNSK--YHYQKAVNHMQASHGEBTFHVLTLTDROGITHKHYVE-PEQEHSPAF 451
Db 475 VLEEMVFRPRPYIYSAMKISTKOQOLYIGSATGVSOQLPHRCQVYGRKACAECCSLARDPY 534
OY 452 NIME-IOPFRRAALIQWMSIDAERRKLIVYSSQWEVSQVHLDEICEVYVGGCGHCCLMSRPY 510
Db 535 CAMDSSCSRIFPLAKRRTRRODIRMGDPLTHGSDLOHNDNSGQLEKIIYYGENSST 594
OY 511 CGMDQGRCIISYSE-RSVL-QSINPAEPHKECPN-PRKDKP--LQ-KYSLA-PNSRY 562
Db 595 FLECSPRKSQALYV-WQFOKQONDDHAYE 621
OY 563 YLSC-PESHRYATYSWR-KHEENVEOSE 588

```

| RESULT | 12 | PRELIMINARY; | PRT; | 754 AA. |
|--|---|--|------|---------|
| ID | 088633 | | | |
| AC | 088633; | | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, CREATED) | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, LAST SEQUENCE UPDATE) | | |
| DT | 01-NOV-1998 | (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | |
| DE | SEMAPHORIN IV ISOFORM A. | | | |
| OS | MUS MUSCULUS (MOUSE). | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA; | | | |
| SC | SCIRIOMGNATHI; MORIDAE; MURINAE; MUS. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | ECKHARDT F., MEYERHANS A.; | | | |
| RT | "Molecular cloning and expression pattern of a murine semaphorin | | | |
| RL | homologous to H-sema IV." | | | |
| DR | SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS. | | | |
| DR | EMBL: AF080091; G3377768; -. | | | |
| SO | SEQUENCE 754 AA; 85133 MW; EC05BDF3 CRC32; | | | |
| Query Match 13.7% Score 682; DB 11; Length 754; | | | | |
| Best Local Similarity 31.8%; Pred. No.1.81e-136; | | | | |
| Matches 141; Conservative 119; Mismatches 140; Indels 43; Gaps 31. | | | | |
| Dh | 171 | PYDPKIDTASALINELINEIAGVYI-DFMGDTAAIFRLGKOTAMRTDQNSRWLNDPSFIH | 229 | |
| QY | 171 | PSPPDENSIVLEFDEGVYSTIRKQEVNGK-IPFRIRIGSESELYTS-DIV-MONPOFIK | 226 | |
| Dh | 230 | AELLIPDSAEERDDKLXFFEFERSAE-APONPAYAVARIGRICINDDGGCCL-VKMSNFI | 287 | |
| QY | 227 | ATIVHQD-QAIDDKIYFFREDNDKRPAPLPLVSVAYOALCGDDGGSSISVSKNMTFL | 285 | |
| Dh | 288 | KARLVCSYGGEDGIETHDELQDVEV-OQTODI-RNPYIAVFTSSGVSFGSAVCYYSM | 345 | |
| QY | 286 | KMLVCS---DAATNKNFNRLQDVFLLPDSGGQMRTIRYGVFSNP-WNY--SAVCYSL | 339 | |
| Dh | 346 | ADIRAVENGPRAHKEGNYOMPPSGMYPAPRGTCPGGTFTPMSTKDYDVEINEMR | 405 | |
| QY | 340 | GIDIDVFTS-SLK-G--YH---SS-LPNRPGRKLPDQ-OP-I-PIETP--OVAD--R | 383 | |

Db 406 THPLMYAVYVLPORRPLVVRT-GAPYRLTTVAVDQVDAADG-RYEVLFGDRTGVQKVI 463
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 384 -HPEVAQVEBPMG--PLKTPLEFSKHYQKAVHMQASHGETFVLLTDRGTHIKVY 440
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 464 VLPDDQVEBELMEVEVEKPEAPVKTMTTSSKQOLYASAVGVTLSLHRCQAYGAA 523
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 441 E-PGEOEHSFAFNIMEIOPFRRAAIOQMSLDAERKRLYVSSQWESQVPLDLCEVYGG 499
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 524 CADCLADPCAMDGQACSRRTASSKRRSRRODVRHGNPIRQCGFNSNANKNAVESVQ 583
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 500 CHGCLMSRDPYCGMDQGRCS-IYSSEK-SVLQSLNPAEPHKECP--NPKPKAPLQKVS 555
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 584 YGVAGSAFLECCQPRSPQATYKW 606
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 556 LA-PNSRYLLSCPMESRHATYSW 577
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||

RESULT 13
 ID 088632 PRELIMINARY; PRT; 785 AA.
 AC 088632;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE SEMAPHORIN IV ISOFORM B.
 DE SEMAPHORIN IV ISOFORM B.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCURIONMATHI; MORIDAE; MORINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ECKHARDT F., MEYERHANS A.;
 RT "Molecular cloning and expression pattern of a murine semaphorin
 RT homologous to H-sema IV".
 RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AF080090; G3377766; -;
 DR EMBL; AF080090; G3377766; -;
 SQ SEQUENCE 785 AA; 88493 MW; 6FFCD8F5 CRC32;

Query Match 13.7%; Score 682; DB 11; Length 785;
 Best Local Similarity 31.8%; Pred. No. 1,816-116;
 Matches 141; Conservative 119; Mismatches 140; Indels 43; Gaps 31;

Db 202 PYDKRLDTASALINEELVAGYI-DEMGTDALFRTLKQATMRTOYNSRWLNDPSEIH 260
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 171 PESPENSLVLEGEDEVSTIRKQYNGK-IPFRRIIGSELYTS--DTV-MQNPQFIK 226
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 261 AELLPDSERDDKLYFFRERSAE-APQNPAYVARIGRICLNDGGHCLL-VKWKSTFL 318
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 227 ATIVHOD-QAYDDKLYFFREDNDKNEAPLNSRAQLRGDGGESSLSVSKWNTFL 285
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 319 KARLVCSVPGEDGJETHEDELQDVV--QOTQDI--RNPIYAVFTSSGSVFERSAVCYASM 376
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 286 KAMLVCS---DAATNKNRNLQDVFLPLDPGSGWRDRTVYGFNSP--WNT--SAVCYSL 339
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 377 ADIMVENGPRAHKEGPNYQWMPFSGKMPYRPGTCGFTTPSMKSTKDYPDDEVINEMR 436
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 340 GDIDKVFRTS-SLK-G-YH---SS-LPNRPCKCLPDQ-QP-I-PIETF--QVAD--R 383
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 437 THPLMYAVYVLPORRPLVVRT-GAPYRLTTVAVDQVDAADG-RYEVLFGDRTGVQKVI 494
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 384 -HPEVAQVEBPMG--PLKTPLEFSKHYQKAVHMQASHGETFVLLTDRGTHIKVY 440
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 495 VLPDDQVEBELMEVEVEKPEAPVKTMTTSSKQOLYASAVGVTLSLHRCQAYGAA 554
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 441 E-PGEOEHSFAFNIMEIOPFRRAAIOQMSLDAERKRLYVSSQWESQVPLDLCEVYGG 499
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 555 CADCLADPCAMDGQACSRRTASSKRRSRRODVRHGNPIRQCGFNSNANKNAVESVQ 614
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 500 CHGCLMSRDPYCGMDQGRCS-IYSSEK-SVLQSLNPAEPHKECP--NPKPKAPLQKVS 555
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 615 YGVAGSAFLECCQPRSPQATYKW 637
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 556 LA-PNSRYLLSCPMESRHATYSW 577
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||

RESULT 14
 ID 013275 PRELIMINARY; PRT; 785 AA.
 AC 013275; 013274;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE SEMAPHORIN IV.
 DE SEMAPHORIN IV.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA NELSON J., BERNALD T.;
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [1]
 RP SEQUENCE OF 394-436 FROM N.A.
 RC TISSUE-PLACENTA;
 RX MEDLINE; 96210603.
 RA SEKIDO Y., BADER S., LATIF F., CHEN J.Y., DUH F.M., WEI M.H.,
 RA ALBANESI J.P., LEE C.C., LERMAN M.I., MINNA J.D.;
 RT "Human semaphorins A(Y) and IV reside in the 3p21.3 small cell lung
 RT cancer deletion region and demonstrate distinct expression
 RT patterns.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 93:4120-4125(1996).
 DR EMBL; AC000063; G1669373; -;
 DR EMBL; U32172; G995788; -;
 DR EMBL; U32171; G995786; -;
 DR PFAM; PF00047; 1g; 1;
 SQ SEQUENCE 785 AA; 88381 MW; CD175765 CRC32;

Query Match 13.4%; Score 681; DB 4; Length 785;
 Best Local Similarity 31.4%; Pred. No. 3,272-116;
 Matches 139; Conservative 122; Mismatches 139; Indels 43; Gaps 31;

Db 202 PYDKRLDTASALINEELVAGYI-DEMGTDALFRTLKQATMRTOYNSRWLNDPSEIH 260
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 171 PESPENSLVLEGEDEVSTIRKQYNGK-IPFRRIIGSELYTS--DTV-MQNPQFIK 226
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 261 AELLPDSERDDKLYFFRERSAE-APQNPAYVARIGRICLNDGGHCLL-VKWKSTFL 318
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 227 ATIVHOD-QAYDDKLYFFREDNDKNEAPLNSRAQLRGDGGESSLSVSKWNTFL 285
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 319 KARLVCSVPGEDGJETHEDELQDVV--QOTQDV--RNPIYAVFTSSGSVFERSAVCYASM 376
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 286 KAMLVCS---DAATNKNRNLQDVFLPLDPGSGWRDRTVYGFNSP--WNT--SAVCYSL 339
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 377 ADIMVENGPRAHKEGPNYQWMPFSGKMPYRPGTCGFTTPSMKSTKDYPDDEVINEMR 436
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 340 GDIDKVFRTS-SLK-G-YH---SS-LPNRPCKCLPDQ-QP-I-PIETF--QVAD--R 383
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 437 SHPLMYAVYVLPORRPLVVRT-GAPYRLTTVAVDQVDAADG-RYEVLFGDRTGVQKVI 494
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 384 -HPEVAQVEBPMG--PLKTPLEFSKHYQKAVHMQASHGETFVLLTDRGTHIKVY 440
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 495 VLPDDQVEBELMEVEVEKPEAPVKTMTTSSKQOLYASAVGVTLSLHRCQAYGAA 554
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 441 EPGEOEHSFA-FNIMEIOPFRRAAIOQMSLDAERKRLYVSSQWESQVPLDLCEVYGG 499
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 555 CADCLADPCAMDGQACSRRTASSKRRSRRODVRHGNPIRQCGFNSNANKNAVESVQ 614
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 500 CHGCLMSRDPYCGMDQGRCS-IYSSEK-SVLQSLNPAEPHKECP--NPKPKAPLQKVS 555
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Db 615 YGVAGSAFLECCQPRSPQATYKW 637
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||
 Oy 556 LA-PNSRYLLSCPMESRHATYSW 577
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : ||

RESULT 15
 ID 015704 PRELIMINARY; PRT; 785 AA.
 AC 015704;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBL:REL. 08, LAST ANNOTATION UPDATE)
DE SEMAPHORIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE; 96226360.
RA ROCHE J., BOLDOS F., ROBINSON M., ROBINSON L., VARELLA-GARCIA L.,
RA SWANTON M., WAGGONER B., FISHEL R., FRANKLIN W., GEMMILL R.,
RA DRABKIN H.;
RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
RT human semaphorin."
RL ONCOGENE 12:1289-1297(1996).
DR EMBL; U33920; G100207; -.
DR PFAM; PF00047; 1g; 1.
SQ SEQUENCE 785 AA; 88385 MW; 9971F14E CRC32;

Query Match 13.7%; Score 679; DB 4; Length 785;

Best Local Similarity 31.6%; Pred. No. 1,07e-135;
Matches 140; Conservative 121; Mismatches 139; Indels 43; Gaps 31;

Db 202 PYDKPLDTASALINELIYAGYI-DEMGTDALIFRTLGQTAMRTDQNSRMLNDPSFIH 260
QY 171 PFSPDENSLVLEFGDEVYSTRKQENYNGK-IPFRIRIGSESELYTS--DTV-MONPOFIX 226
Db 261 AELIPDSAEKNDKLYFFRERSAE-AQSPAYVARIGRICLNDGHCCL-VNKMSTFL 318
QY 227 ATIVHOD-QAYDDKIYFFREDNPKNEAPLNSRVADLCRGDGGESSLSVSKMTEFL 285
Db 319 KARLYCSVPGEDGIEHDELDQYFV-QOTQDV--RNPVIYAVFTSSGSVFERSAVCYISM 376
QY 286 KAMLYCS--DAITKNKNRLQDVFLPDPGQMRDTRVYGVFSNP-WNY--SAVCYISL 339
Db 377 ADIRWVNGPFAHKEGPNYQMMPEFGKMPYRPGTCPGGFTFPKSKSTKDYPDEVINEMR 436
QY 340 GDIDKVFRTS-SLK-G-YH---SS-LPNRPCKCLPDQ-QP-I-PTEFE-QVAD--R 383
Db 437 SHPLMYQAVYPLQRRPLVVR-GAPYRLTTAVDOVDAGDG-RYEVFLGTDRTGVOKVI 494
QY 384 -HPEVAQKVEPMG--PLKTPLEHSHKYQKAVAHMQASHGETFHVLYLTDRGTIRKVV 440
Db 495 VLPKDDQEMELMLEEVVFQDPAVPKTMITSKROQLYVASAVGVTHLSLHCOAYGAA 554
QY 441 E-PGEOHSEFANIMEIOPFRRAAIQTMISLDAERKLIYSSQWFEVSQVPLDLCEVIYGG 499
Db 555 CADCCCLARDPYCAMDQACSRYTASSKRRSRQDVRHGNPIROC GFNSNANKNAVESVQ 614
QY 500 CHGCLMSRDPCGMDQGRGIS-IYSSEK-SVLQSLNPAEPHKECP--NPKPDKAPLQKVS 555
Db 615 YGVAGSAFLLECOPRSPQATYKW 637
QY 556 LA-PNSRYLLSCPMESRHATYSW 577

Search completed: Tue Aug 3 16:32:25 1999
Job time : 127 secs.